

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12Q 1/68, C12P 19/34 // C12N 15/09, 15/10, 15/66, C07H 21/04</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 96/41005</b> <b>(43) International Publication Date:</b> 19 December 1996 (19.12.96)
<b>(21) International Application Number:</b> PCT/US96/09208 <b>(22) International Filing Date:</b> 6 June 1996 (06.06.96) <b>(30) Priority Data:</b> 08/479,723 7 June 1995 (07.06.95) US <b>(71)(72) Applicants and Inventors:</b> MURTAGH, James, J. [US/US]; 511 Calibre Woods Drive, Atlanta, GA 30329 (US). THUNNISSEN, Frederik, B., J., M. [NL/NL]; Pastoor Kribsweg 21, NL-6212 BS Maastricht (NL). <b>(74) Agents:</b> PERRYMAN, David, G. et al.; Needle & Rosenberg, Suite 1200, 127 Peachtree Street, N.E., Atlanta, GA 30303-1811 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> METHODS FOR NUCLEIC ACID DETECTION, SEQUENCING, AND CLONING USING EXONUCLEASE		
<b>(57) Abstract</b> <p>The present invention provides a method of detecting the presence of a nucleotide sequence within a double-stranded DNA in a sample comprising: (a) digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA; (b) hybridizing the single-stranded DNA with i) a first nucleic acid probe adapted with a moiety which can be captured to a solid support, and ii) a second nucleic acid probe labeled with a detectable moiety which can hybridize with the single-stranded DNA adjacent the hybridized first nucleic acid probe; (c) ligating the hybridized first and second nucleic acid probes; (d) capturing the moiety on the first nucleic acid probe hybridized to the DNA to the solid support; (e) denaturing the ligated first and second nucleic acid probes from the hybridized single-stranded DNA; (f) removing uncaptured labeled probe; and, (g) detecting the presence of captured detectable moiety, the presence of captured detectable moiety indicating the presence of the nucleotide sequence within the double-stranded DNA in the sample.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

## METHODS FOR NUCLEIC ACID DETECTION, SEQUENCING, AND CLONING USING EXONUCLEASE

### Background of the Invention

5

DNA analytic and hybridization methods to specifically detect and identify small amounts of nucleic acids are now indispensable to molecular biology, the genetic engineering industry, and molecular medicine. Of these, procedures that amplify target nucleic acid sequences—such as the polymerase chain reaction (PCR)<sup>1</sup>, ligase  
10 amplification, and transcription-based systems derived from PCR—have proven most effective<sup>2</sup>.

Difficulties in critical analytic steps following DNA amplification, however, have prevented the technology from being more widely applied in clinical laboratories  
15 and the biotechnology industry. Analysis of DNA products following amplification usually takes one of three forms: (1) colorimetric hybridization analysis, (2) cloning, (3) direct sequencing<sup>1-3</sup>. Each of these procedures has proved more difficult than initially anticipated, in part because of special properties of amplified DNA that differ from nonamplified DNA. Numerous investigators have documented these well-known  
20 difficulties in analyzing amplified DNA, resulting from the following:

1. Complementary amplified DNA strands are usually uneven (or "ragged") because some strands are usually incompletely extended, and extraneous nucleotides are often added to 3' ends. "Ragged" DNA ends cannot be ligated to  
25 blunt end vectors, and is inefficiently bound by modifying enzymes<sup>1-5</sup>.
2. Independently of the "raggedness" of DNA ends, base pairing at extreme DNA ends is believed to be unstable (an effect referred to as "breathing" of DNA ends). This effect appears to lead to increased susceptibility to exonuclease, and  
30 further difficulties in binding modifying enzymes, including restriction endonucleases. PCR-amplified DNA is widely found to cut inefficiently by

many restriction endonucleases—including NotI, XbaI, XhoI, and SmaI—preventing analysis and cloning by many strategies<sup>5</sup>.

3. Incompletely used substrates and reaction by products of amplification procedures can interfere with subsequent analysis. Interfering reactants include nucleotides and oligonucleotides; interfering byproducts include pyrophosphates.

4. Extraneous DNA, including "primer dimer," is often coamplified with target, confusing subsequent analysis and competing with target in cloning and sequencing steps. For this reason, amplified DNA must be extensively purified prior to these analytic procedures.

5. Amplified DNA strands tend to rapidly reanneal and exclude hybridized probes<sup>1-3,6</sup>. This effect has been noted to create special difficulties in the sequencing of double-stranded PCR products. Numerous strategies have been suggested for converting PCR fragments into single strands so that they can be more easily analyzed by hybridization, sequencing, and cloning, including asymmetric PCR and magnetic beads. However, none has proved entirely satisfactory.

Many strategies to circumvent difficulties in analyzing amplified DNA require oligonucleotide primers that have 5' modifications, such as chemical groups including biotin, digoxigenin and fluorescein that can aid in detection and/or purification of amplified DNA.

Other strategies require sequence additions, such as 40 nt GC clamps, RNA promoters, restriction endonucleases etc. Since the needs of individual experiments vary greatly, investigators often need several oligonucleotides of similar or identical 3' sequence targeted to particular genes under study, but have various different 5' modifications. The current invention aims in large part to simplify and unify many



different protocols requiring different 5' additions to amplified DNA. Defined 6-12 base sequences are added to the 5' of each amplification primer. After resulting amplification, three defined 6-12 base sequences are used to attach universal zipper adapter primers, so that useful desired chemical groups and functional sequences can be easily added to amplified DNA. Thus, the invention is a unified, integrated system to create an "economy of primers," so that an individual amplification product can be used in many different capacities without resynthesizing amplification primers.

The invention also simplifies detection and characterization of amplified DNA. Previously, colorimetric PCR hybridization assays have been investigated as an alternative to gel electrophoresis and Southern blot analysis of PCR-amplified DNA. Such gels and Southern blots have proven too time-consuming and tedious for typical clinical laboratories<sup>8</sup>. Solid-phase colorimetric PCR assays capture denatured amplification products on probes bound to nylon membranes (as in "reverse dot blots"<sup>6-7</sup>) or to microtiter plates<sup>2,3</sup>. However, the sensitivity of these assays suffers due to the tendency of the denatured PCR product strands to reassociate and exclude oligonucleotide probes, and steric interference between the bound oligonucleotides and the solid support, which impedes hybridization to nucleic acids in solution<sup>5,6</sup>. In some cases, colorimetric detection is improved by creating single-strand PCR products through asymmetric PCR that can associate with bound probes without interference<sup>7,9</sup>. Unfortunately, asymmetric PCR is notoriously difficult to reproduce, and does not lend itself to automation. Reamplification of the PCR product using internal or "nested" primers may improve assay sensitivity, but is costly, and compounds DNA contamination problems in clinical laboratories with concomitant false positive results.

Numerous different cloning procedures for analyzing amplified DNA can be found in the art, but all remain inefficient, expensive, and tedious. Restriction endonucleases do not cut PCR produced DNA well, and untreated PCR products cannot be cloned into blunt-end vectors unless the ragged ends of PCR are first repaired. Blunt end ligation is inefficient under any circumstances.<sup>4</sup>

"TA" cloning vectors exploit the 5' adenosine residues that are sometimes added to the 3' of PCR products to clone into a vector with thymidine residues (Invitrogen, San Diego, CA). Major problems with the use of these vectors includes instability of the terminal thymidine residue, inefficient transformation, and the limitation of using  
5 Taq polymerase to generate the terminal adenosine residues on the PCR product. Hybridization of single-base overhangs is inefficient, and these vectors do not work well in most laboratories.

Methods using T4 DNA polymerase to clone PCR products have recently been  
10 introduced into the art. These methods seek to introduce extended cohesive ends into PCR products complementary to mirror cohesive ends placed through recombinant methods into vectors. Stoker et al. performed PCR amplification with two primers which are homologous to the cohesive termini created by *AccI* and *XmaI*, respectively. The PCR products are treated with T4 DNA polymerase to remove 3' terminal  
15 sequences. After heat inactivation of the polymerase, the products are ligated to plasmid cut with *AccI* and *XmaI*<sup>10</sup>. Aslanidis et al. generated clonable PCR fragments with 5' ends containing an additional 12 nucleotide sequence which lacks dCMP<sup>11</sup>. After amplification, products were digested by T4 DNA polymerase in the presence of dGTP; the fragments thus have 5'-extending single-stranded tails of a defined sequence  
20 and length. In the same way, the plasmid vector was amplified with primers homologous to sequences in the multiple cloning site. The vector oligos have additional 12 nucleotide tails complementary to the tails used for fragment amplification, permitting the creation of single-stranded ends with T4 DNA polymerase in the presence of dCTP<sup>11</sup>. This was similar to the method of Kuijper et al. ("prime"  
25 cloning) who introduced sequences into plasmids and lambda phage deficient in the nucleotide dTTP<sup>12</sup>. After digestion with restriction endonucleases and T4 DNA polymerase in the presence of dTTP, these vectors accept PCR fragments with mirror cohesive ends. Other investigators have used this technique with variable success<sup>13-15</sup>.

30 Each of the methods for cloning using T4 DNA polymerase suffers from severe difficulties. First, sequences to be made cohesive must be specially engineered into the

vector either by PCR<sup>11</sup> or via gene construction<sup>116</sup>. These methods are therefore unsuitable for use with most commonly used vectors. Second, PCR products must be extensively purified prior to cloning. The 3' exonuclease of T4 DNA polymerase is active only in the absence of 3 of 4 nucleotide species. Third, T4 DNA polymerase is found to be quite labile, and most commercially available lots of this enzyme are found to be unsuitable for this purpose. Users of these methods have documented the need to test multiple enzyme lots prior to identifying suitable T4 enzymes.

Kaling et al. replaced T4 DNA polymerase with exonuclease III as an enzyme for creating cohesive ends in PCR cloning<sup>5</sup>. However, this procedure requires kinased primers, and generates very limited cohesive ends (4 bases) corresponding to 5' protruding restriction endonuclease sites in plasmid. This method is therefore restricted to plasmids with appropriate restriction endonuclease sites. In addition, exonuclease III is often found to contain single-strand nuclease activity that can digest away cohesive protruding single-strand ends and lower cloning efficiency.

Cloning procedure using uracil DNA glycosylase (UDG)<sup>17</sup> are found in the art<sup>18-20</sup>. However, these procedures apply only to specially constructed vectors containing uracil (dUTP). Specially constructed PCR primers made with uracil phosphoramidite are also required, material that is expensive and unavailable to many laboratories.

Analysis of amplified DNA sequence can also be accomplished by direct sequencing without cloning to vectors. However, most laboratories have found sequencing of double-stranded PCR product to be inefficient and unsuitable for routine use. Several methods for converting PCR fragments into single strands, including the asymmetric PCR protocol of Gyllensten and Erlich<sup>9</sup>, the affinity strand separation method of Mitchell and Merrill<sup>21</sup>, and the magnetic strand separation method of Uhlén and coworkers<sup>22</sup>. However, the efficiency of asymmetric PCR is notoriously variable from sample to sample, and the other methods require expensive materials including biotinylated primers and streptavidin-coated beads.

Consequently, there remain several needs in the art for DNA analytic procedures to specifically detect, identify, and manipulate amplified DNA. In DNA diagnostics, the greatest need is a rapid, economical, highly sensitive colorimetric DNA hybridization test to specifically detect amplification products without running electrophoresis gels or performing Southern blot experiments. Preferably, such a test could be performed in a format familiar to clinical laboratories, such as a microtiter plate, with greater sensitivity than current tests to detect denatured DNA. In many cloning experiments, a need exists for a convenient rapid procedure to flexibly recombine PCR-generated DNA into vector plasmids and bacteriophage with high efficiency, without the need to alter vector by addition of sequence, or alteration with restriction endonuclease. In other experiments, it is also important that vectors can be altered for rapid transfer of adapted DNA. The procedure needs to bypass the need for T4 DNA polymerase, an enzyme of variable efficiency and stability. PCR products should not require extensive purification. In sequencing, a need exists for a rapid method to convert amplified DNA partly or wholly into single strands without expensive or time-consuming protocols, and to extend the number of base sequences obtained per sequencing reaction. The present invention provides such methods as well as many related advantages.

20

## SUMMARY OF THE INVENTION

The present invention provides a method of detecting the presence of a nucleotide sequence within a double-stranded DNA in a sample comprising: a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA, b. binding the single-stranded DNA with a nucleic acid probe which selectively hybridizes with the single-stranded DNA, and c. detecting hybridization between the single-stranded DNA and the nucleic acid probe, the existence of hybridization indicating the presence of the nucleotide sequence within the double-stranded DNA in the sample. The present

30

invention further provides a method of detecting the presence of a nucleotide sequence in a sample comprising DNA which is the product of a DNA amplification technique.

The invention also discloses a method of detecting the presence of a nucleotide  
5 sequence in a sample comprising: a. amplifying the nucleotide sequence using a first primer containing deoxyuridine monophosphate and a second primer not containing deoxyuridine monophosphate to yield a double-stranded DNA having a polyuridylated 5' end on one strand, b. partially disassociating the resultant double-stranded DNA with uracil DNA glycosylase to form a 3' overhang on the polyuridylated 5' end, c. digesting  
10 the double-stranded DNA with a 3' exonuclease which digests only a 3' strand from an end opposite the polyuridylated 5' end for a time sufficient to convert at least a portion of the double-stranded DNA to a single-stranded DNA, d. binding the single-stranded DNA with a nucleic acid probe which selectively hybridizes with the single-stranded DNA, and e. detecting hybridization between the single-stranded DNA and the nucleic  
15 acid probe, the existence of hybridization indicating the presence of the nucleotide sequence in the sample.

The present invention also provides a method of adapting a target DNA for subsequent manipulation comprising: a. annealing to the target DNA a first primer  
20 having a 3' terminal region homologous to a portion of the target DNA, and a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization, b. amplifying the annealed target DNA and primer to yield double-stranded DNA, c. partially digesting the double-stranded DNA with a 5' exonuclease for a time sufficient to convert only a portion of the double-stranded DNA  
25 to single-stranded DNA having a 3' terminal sequence, d. selectively hybridizing a second primer with a portion of the 3' terminal sequence of the single-stranded DNA, wherein the second primer contains a functional group capable of subsequent manipulation, and e. contacting the partially digested primer-DNA hybrid with a DNA polymerase to form a contiguous double-stranded DNA containing the functional  
30 group, thereby adapting the double-stranded DNA for subsequent manipulation.

It is also a purpose of this invention to provide a method of adapting a nucleic acid for sequencing comprising: a. amplifying the nucleic acid using a first primer containing deoxyuridine monophosphate and a second primer not containing deoxyuridine monophosphate to yield a double-stranded DNA with a polyuridylated 5' end on one strand, b. partially disassociating the resultant double-stranded DNA with uracil DNA glycosylase to form a 3' overhang on the polyuridylated 5' end, c. digesting the double-stranded DNA with an 3' exonuclease which digests only the 3' strand from the end opposite the polyuridylated 5' end for a time sufficient to convert at least a portion of the double-stranded DNA to a single-stranded DNA, d. extending the partially digested 3' strand with a polymerase such that sequencing by a dideoxynucleotide process can be performed on the single-stranded DNA.

Furthermore, the invention provides a method of adapting a double-stranded DNA for insertion into a vector comprising digesting the double-stranded DNA in with a non-polymerasing DNA exonuclease for a time sufficient to create cohesive single-stranded DNA terminal regions which will hybridize with a vector having complementary cohesive single-stranded regions.

The present invention also provides a method of inserting a target DNA into a target vector for cloning comprising: a. annealing to the target DNA a first primer having a 3' terminal region homologous to a portion of the target DNA, and a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the denatured DNA a second primer having a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, and having a 5' terminal sequence homologous to single-stranded regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the

cohesive single-stranded insert DNA regions with the target vector having homologous single-stranded regions to create a circular gapped vector-insert DNA hybrid, and h. transfecting the circular gapped vector-insert DNA hybrid into a host cell.

5           Furthermore, the invention provides a method of adapting a target vector for subsequent recombination comprising: a. annealing to a target DNA a first primer having (1) a 3' terminal region homologous to a portion of the target DNA (2) a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization and (3) a sequence encoding a restriction endonuclease site  
10   located at the junction of the 3' terminal and the 5' terminal region of the primer, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the denatured DNA a second primer having a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, and having a 5' terminal sequence homologous to single-stranded  
15   regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the cohesive single-  
20   stranded insert DNA regions with the target vector having homologous single-stranded regions to yield a circular plasmid, h. transfecting the plasmid into a host cell, i. culturing the host cell containing the plasmid, j. purifying the plasmid from a lysate of host cell culture, k. contacting the plasmid with a restriction endonuclease which cleaves the site originating from the first primer such that the DNA originating from the  
25   first primer remains on the target vector, and l. exposing the target vector to an exonuclease for a time sufficient to create cohesive single-stranded DNA regions originating from the first primer on the vector, thereby adapting the vector for subsequent recombination.

30           The invention provides a method of adapting an insert DNA for subsequent recombination with a vector comprising: a. annealing to a target DNA a first primer

having a 3' terminal region homologous to a portion of the target DNA, and a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the

5 denatured DNA a second primer having (1) a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, (2) a 5' terminal sequence homologous to single-stranded regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, and (3) a sequence encoding a restriction

10 endonuclease site located at the junction of the 3' terminal sequence and the 5' terminal sequence of the second primer, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the cohesive single-

15 stranded insert DNA regions with the target vector having homologous single-stranded regions to yield a plasmid, h. transfecting the plasmid into a host cell, i. culturing the host cell containing the plasmid, j. purifying the plasmid from a lysate of host cell culture, k. contacting the plasmid with a restriction endonuclease which cleaves the site originating from the first primer such that the DNA originating from the first primer remains on the insert DNA, and l. exposing the target vector to an exonuclease for a

20 time sufficient to create cohesive single-stranded DNA regions originating from the first primer on the vector, thereby adapting the insert DNA for subsequent recombination with a vector.

The invention provides a method of detecting the presence of a nucleotide

25 sequence within a double-stranded DNA in a sample comprising: a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA; b. hybridizing the single-stranded DNA with i) a first nucleic acid probe adapted with a moiety which can be captured to a solid support, and ii) a second nucleic acid probe labeled with a detectable moiety which

30 hybridizes with the single-stranded DNA adjacent the hybridized first nucleic acid probe; c. ligating the hybridized first and second nucleic acid probes; d. capturing the



moiety on the first nucleic acid probe hybridized to the DNA to the solid support; e. denaturing the ligated first and second nucleic acid probes from the hybridized single stranded DNA; f. removing uncaptured labeled probe; and g. detecting the presence of captured detectable moiety, the presence of captured detectable moiety indicating the presence of the nucleotide sequence within the double-stranded DNA in the sample.

The invention provides a kit comprising: a) an exonuclease; b) a first probe which binds to target DNA which is adapted with a capture moiety; c) a second probe which binds to target DNA which is labeled with a detectable moiety; and d) a ligase.

## BRIEF DESCRIPTION OF THE DRAWINGS

**Figure 1** illustrates and defines the structures and terms used in many aspects of the invention. **A.** Zipper-containing gene-specific primers (zgsp) are constructed with 20-24 bases complementary to target to be amplified. In some cases, the target-specific domain is followed by an *Apa* I restriction site, which may be used to strike target-specific sequence from additions. The 5' end of the primer consists of 6-12 bases defined as "zipper sequences" or "zippers." All "forward" or "upstream" gene-specific primers (defined as the PCR primer synthesized in the "sense" direction to the most 5' region of target DNA) are synthesized with 6-12 bases of the following sequence, defined as the "forward zipper": CGAGGGAAGAGG (SEQ ID NO:1). Each "reverse" or "downstream" gene-specific primer (defined as the PCR primer synthesized in the antisense orientation to target, and complementary to the 3' end of target) are synthesized with a different 6-12 bases of sequence, defined as the "reverse zipper": CGCACGCGGGAG (SEQ ID NO:2). Underlined sequence was found to be minimal nucleotide additions for the conversion. **B.** Zipper adapter primers (ZAPs) are constructed with 12-base zipper sequences at the 3'; thus, forward ZAPs have the sequence CGAGGGAAGAGG (SEQ ID NO:1) placed at the 3'; reverse ZAPs have the sequence CGCACGCGGGAG (SEQ ID NO:2) at the 3'. Following zipper sequences, most ZAPs are synthesized with a *Pvu* I restriction site, to permit cutting of insert from

additions so that zipper sequence remains with insert. At the 5' of ZAPs are sequences including vector-specific sequences for cloning, T7 RNA promoters for transcription, sequence initiation sites, etc., on chemically modified nucleotides containing biotin, fluorescein, etc.

5

**Figure 2** illustrates an exonuclease-amplification coupled confirmation technique (EXACCT), in format A. **A.** DNA is digested partially or to completion using exonuclease. Direction proceeds from the end of both strands of double-stranded DNA in standard buffers used for amplification. **B.** Digestion is complete when  
10 exonuclease digestion meets near the middle of the PCR product; the exonuclease is unable to alter single-stranded nucleic acids, so digestion is terminated. **C.** Two separate oligonucleotides complementary to sequence in the first third of the PCR product are solution-hybridized to digested PCR product. One oligonucleotide is biotin-labeled (●) and serves to bind the complex to a solid support. The other  
15 nucleotide bears a detector group (\*) such as digoxigenin, fluorescein, or <sup>32</sup>P moiety, etc. The method can be used in detecting amplified DNA, segments of restricted repetitive DNA, as from mitochondria, and in capturing and detecting highly repetitive ribosomal rRNA as in a detection scheme for mycobacteria.

**Figure 3** illustrates a second exonuclease-amplification coupled DNA diagnostic scheme in format B. In this invention, one primer used to amplify DNA bears a polyuridine (U) tail. **A.** Treatment of DNA with UDG creates a 3' overhang, protecting that DNA end from digestion with exonuclease III. **B.** Exonuclease III digestion proceeds from nonprotected end. **C.** Complete digestion of the uridine-  
25 containing strand leaves a complementary full-length single-strand DNA **D.** Two separate oligonucleotides are hybridized to the protected single-strand DNA, similar to Fig. 1. Other strategies to selectively protect one end of the DNA molecule from exonuclease digestion can also be used, such as introduction of phosphothioate linkages into 5' or 3' ends of oligonucleotides.

30

**Figure 4** illustrates a general ligation using exonuclease extension (GLUEE).

**A.** DNA is amplified by zipper-containing gene specific primers (zgsp) PCR primers containing zippers (6-12 base defined sequences) at the 5'. **B.** PCR product is treated briefly with a 5' exonuclease, such as T7 gene 6 to create cohesive ends that span the zipper sequence. **C.** Adapter primers targeted to zipper sequence are annealed at room temperature. **D.** Extension of adapter primers on zipper-cohesive ends produce double-strand DNA containing functional groups introduced through adapters. A small amount of adapters can be used to introduce desired sequences into innumerable different amplification products.

**Figure 5** illustrates a pair of zipper adapter primers (ZAPs) that have been proved particularly useful. **A.** ZAP-F is a 66-base oligonucleotide (SEQ ID NO:16) that can be attached to the forward (upstream) end of any zipper-containing DNA. Its sequence contains useful functional groups including an M13 forward universal sequencing primer (especially useful in dye-labeled DNA sequencing reactions in automated fluorescent sequencers), a T7 RNA promoter, EcoRI and Pvu 1 restriction sites. **B.** ZAP-R is an 81-base oligonucleotide (SEQ ID NO:17) containing an M13 reverse sequencing site, SP6 RNA promoter, HindIII and Pvu 1 restriction site. Sequences corresponding to zipper sequences, to which the adapters are targeted in GLUEE reactions or secondary PCR, are underlined. These adapters have been synthesized with and without biotin, so that either end can be easily attached to solid support. After these adapters have been attached to DNA, either strand can be transcribed into cRNA in either liquid or solid phase. DNA inserts can be sequenced by various mechanisms. The sequences of ZAP-F and ZAP-R can also be complementary to sequence at the insertion site of pGEM-3Z, so that sequences can be efficiently cloned using procedures shown in Figure 8.

**Figure 6** illustrates joining of zipper-cohesive inserts to solid support. **A.** Oligonucleotides, (SEQ ID NO:1) linked through 5' attachments to solid support (such as magnetic beads, polystyrene plastic, sephadex beads, microtiter plates, nylon membrane, glass, etc.), are added to DNA with zipper cohesive ends produced through

exonuclease treatment. **B.** The 3' ends contain zipper sequence and anneal to the cohesive DNA ends containing zipper sequence. **C.** Extension by polymerase results in double-strand, fully-extended DNA attached to solid support. This DNA can be used in sequencing reactions, as are DNA affinity columns to capture related nucleic acids, or in  
5 radioactive and nonradioactive nucleic acid detection. The same procedure illustrated may be used to recombine DNA with numerous other functional groups.

**Figure 7** illustrates introduction of adapters into DNA through secondary PCR. The same zipper sequences and adapters are used as in Fig. 5, except that adapters are  
10 used at high concentration to amplify zipper-containing DNA in a secondary (or "piggyback") PCR. **A.** First PCR primed by zipper-containing gene-specific primer (zgsp) having zipper sequences at 5' (see Fig. 1.A.) yields double-strand DNA. **B.** Secondary PCR ("piggyback" PCR) primed by ZAP (Fig 1.B.) targeted to zipper domains introduced during first PCR. The result is double-strand DNA containing the  
15 sequences and functional groups introduced through the adapter domain of ZAP.

**Figure 8** illustrates cloning of DNA into plasmid vector by the invention. **A.** Sequence complementary to vector is affixed to zipper-containing DNA by GLUEE (Fig. 4) or secondary PCR (Fig. 7). **B.** Resulting DNA is briefly treated with  
20 exonuclease (5' or 3' digesting) to expose cohesive ends. **C.** Cohesive DNA is annealed to complementary cohesive ends of vector that has been similarly treated with exonuclease. Annealed DNA is transfected into host cells (bacteria or eukaryotes).

**Figure 9** illustrates construction of pZGEM (SEQ NOS: 63 and 64) by  
25 conventional cloning procedures.

**Figure 10** illustrates creation of a vector with zipper-cohesive ends. Recombinant plasmids with zipper-bearing inserts (such as from Fig. 5) are engineered to contain restriction sites (in the example, the restriction site Apa 1 has been used) at  
30 the target with boundary (arrow). **A.** Digestion of plasmid with Apa 1 followed by **B.** exonuclease treatment generates a vector with zipper-sequence that can anneal to any

complementary cohesive zipper-containing DNA containing ends. **B.** Exonuclease treatment yields cohesive ends.

**Figure 11** illustrates regeneration of inserts containing zipper-cohesive ends from plasmids without further PCR. In our example, PCR primers and adapters were engineered to create Pvu 1 sites at the zipper-vector boundary. After digestion with **A.** Pvu 1 and **B.** exonuclease, cohesive zippers are again created, so that insert can be shuttled to any zipper-containing vector (as created in Fig. 10). Cohesive inserts may also be joined to additional adapters by GLUEE (Fig. 4) or by secondary PCR (Fig. 7), then cloned into vectors that have not been previously altered with zipper sequence.

**Figure 12** illustrates utility of zipper adapters as an adjunct in cloning into vectors containing uridine. Procedures in the art exist to clone via digestion of uridines by uridine DNA glycosylase (UDG). However, such procedures require many PCR primers to be made with uridine, an expensive proposition that is not feasible for many laboratories. Using the invention, universal adapters containing uridines are attached to zipper-containing DNA (using methods of Fig. 3). Resulting DNA can be treated by UDG and cloned. **A.** Primary amplification of target using zipper containing gene-specific primers (zgsp). **B.** Secondary amplification using zipper adapter primers (ZAPS) containing uridines (U-FOR and U-REV). **C.** Digestion of double-strand PCR product with UDG to create cohesive ends. **D.** Cloning of PCR product into complementary cohesive ends of pSPORT vector (BRL, Bethesda, MD).

**Figure 13** illustrates introduction of zipper sequences into DNA that have not been amplified. Although the current invention is primarily intended as a method to manipulate and clone DNA produced by amplification, zippers and adapters can be introduced into other nonamplified DNAs using ligase. Zipper sequences can also be introduced at the 5' of oligonucleotides used to prime mRNA via reverse transcriptase into cDNA. Second-strand cDNAs can be joined to a second zipper by tailing or ligation. The examples of Fig. 13 are described by SEQ ID NOS: 65-74.

**Figure 14** illustrates sequencing using the present invention with a combination of UDG and ExoIII enzymes. **A.** 3' overhangs are produced in DNA that contains uridines at one end only (as in Fig. 3). **B.** Exonuclease III digestion for different time points creates variable-length 5' overhangs. **C.** After exonuclease III digestion is terminated by heating, sequencing can commence from the variable-length strand.

**Figure 15** illustrates sequencing by the method illustrated in Figure 14 when one strand has been completely digested. **A.** 3' overhangs are produced in DNA that contains uridines at one end only (as in Fig. 3). **B.** Exonuclease III digestion for different time points creates variable-length 5' overhangs. **C.** After complete digestion of the susceptible DNA strand, the protected DNA strand consists of a full-length single strand. **D.** Sequencing primer is annealed to the remaining strand so that dideoxy sequencing may commence.

**Figure 16** illustrates detection by the Point-EXACCT method. In this example, Point-EXACCT is shown for cell line Calu-1, heterozygous for K-ras codon 12 base 1, containing both wild type guanine (G) and mutation thymidine (T). DNA from the cell line was amplified using primers of the first exon of c-Ki-ras gene outside the codon 12 region under conditions previously described. Following amplification, the reaction products digested with T7 gene 6 exonuclease. Subsequently, an aliquot of digested product was added to hybridization buffer, a digoxigenin-labeled probe and a biotinylated probe containing the point mutation at the 3' side. Four different biotinylated probes containing the wild type or point mutation sequence were separately used for hybridization. Following hybridization, the products were transferred to a 96-well streptavidin-coated flat-bottomed microtiter plate (coupling). After coupling and washing, the products were incubated with T4 DNA ligase. In case of perfect complementarity, the enzyme T4 DNA ligase covalently joins the 5' biotinylated probe and the 3' detection probe. If the probes and target are mismatched at their junction, a covalent bound is not formed. After ligation, denaturation (to remove the non-ligated Dig labeled probes) and subsequent washing, remaining products are detected by HRP-conjugated anti-digoxigenin-antibody.

**Figure 17** illustrates the relative optical density of the Point-EXACCT procedure obtained after ligation for five different cell lines: HL60, A549, Calu-1, SW480 and SW1116. Optical density from the ligated products is expressed relative to the (maximal) optical density of the controls: a hybridized sample without ligation (i.e., EXACCT only). The biotinylated probes mismatching to the known base composition for Ki-ras codon 12 in cell lines tested are regarded as negative controls.

**Figure 18A & 18B** illustrate reconstruction experiments after serial dilutions of vital for Ki-ras codon 12 mutant and wild type cell line and testing with Point-EXACCT. The results of the reconstruction experiment are shown in **Figure 18A** for mutant cell line A549 and in **Figure 18B** for mutant cell line Calu-1. In these figures, the optical densities obtained from the different dilutions of the mutant cell line in the wild type cell line HL60 were expressed relative to the (maximal) optical density of the undiluted products. 3 x standard deviation of the mean of the background values (obtained after incubation with the biotinylated probes mismatching to the cell line tested and represented by the broken line) equals 10.0 and 8.1 relative optical density units for A549 and Calu-1, respectively.

20

## DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a method of detecting the presence of a nucleotide sequence within a double-stranded DNA in a sample comprising: a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA, b. binding the single-stranded DNA with a nucleic acid probe which selectively hybridizes with the single-stranded DNA, and c. detecting hybridization between the single-stranded DNA and the nucleic acid probe, the existence of hybridization indicating the presence of the nucleotide sequence within the double-stranded DNA in the sample. An example of this method is shown in Fig. 2.

By "selective hybridization" it is meant that hybridization only substantially occurs with the targeted nucleic acid. Thus, through "selective hybridization" one can detect the presence of a target nucleic acid in an unpure sample. The preferred length of the nucleic acid probes is about 20 bases with an optimum functional range from about 12 to 100 bases. Various exonucleases may be selected for use in this method including T7 gene 6, exonuclease III, T4 DNA polymerase, non-recombinant T7 DNA polymerase, and lambda exonuclease.

The nucleic acid probes utilized in these methods may be labeled with a detectable moiety, and hybridization determined by detecting the presence of the detectable moiety. The single-stranded DNA may also be bound with at least two selectively hybridizing nucleic acid probes which bind unique regions on the single-stranded DNA. Additionally, a first nucleic acid probe can be labeled with biotin and a second nucleic acid probe can be labeled with a detectable moiety, and hybridization determined by binding the biotin on a streptavidin-coated solid support, removing unbound DNA and second probe, and detecting the presence of the moiety of the second probe which is bound to the DNA. The nucleic acid probes may be bound to solid support, such as biotin, either before or after hybridization through either covalent or non-covalent attachment of probe to solid support. The detectable moiety can be selected from a wide range of detectable substances including digoxigenin, fluorescein, acridine-ester, radioactive isotopes, or enzymes.

Depending on the location of desired probe hybridization the double-stranded DNA can be contacted with an exonuclease for a time sufficient to convert the entire double-stranded DNA to single-stranded DNA. The amount of time required will therefore vary from about 30 seconds to over 10 minutes or more, depending on the length of double-stranded DNA necessary to convert to single strands.

The present invention described above can likewise be applied to a method of detecting the presence of a nucleotide sequence which is the product of a DNA amplification technique. Various amplification techniques which are contemplated



include, for example, polymerase chain reaction, ligase chain reaction, transcription amplification system, isothermal chain reaction based on RNA transcription.

The invention also discloses a method of detecting the presence of a nucleotide  
5 sequence in a sample comprising: a. amplifying the nucleotide sequence using a first  
primer containing deoxyuridine monophosphate and a second primer not containing  
deoxyuridine monophosphate to yield a double-stranded DNA having a polyuridylated  
5' end on one strand, b. partially disassociating the resultant double-stranded DNA with  
uracil DNA glycosylase to form a 3' overhang on the polyuridylated 5' end, c. digesting  
10 the double-stranded DNA with a 3' exonuclease which digests only a 3' strand from an  
end opposite the polyuridylated 5' end for a time sufficient to convert at least a portion  
of the double-stranded DNA to a single-stranded DNA, d. binding the single-stranded  
DNA with a nucleic acid probe which selectively hybridizes with the single-stranded  
DNA, and e. detecting hybridization between the single-stranded DNA and the nucleic  
15 acid probe, the existence of hybridization indicating the presence of the nucleotide  
sequence in the sample. Fig. 3 provides an example of this method.

The preferred number of uracils to be attached to the polyuridylated 5' end is  
between 4 and 6 but should be at least 3. A particularly useful exonuclease is  
20 exonuclease III. The time recommended for digestion is between 1 and 10 minutes, but  
may be between 30 seconds and 30 minutes. The double-stranded DNA may be  
contacted with exonuclease for a time sufficient to convert the entire double-stranded  
DNA to single-stranded DNA. Depending upon the length of the DNA and the location  
of the primer hybridization region, this time may be greater than 10 minutes.

25

The present invention also provides a method of adapting a target DNA for  
subsequent manipulation comprising: a. annealing to the target DNA a first primer  
having a 3' terminal region homologous to a portion of the target DNA, and a 5'  
terminal region not homologous to the target DNA and of a length sufficient for future  
30 selective hybridization, b. amplifying the annealed target DNA and primer to yield  
double-stranded DNA, c. partially digesting the double-stranded DNA with a 5'

exonuclease for a time sufficient to convert only a portion of the double-stranded DNA to single-stranded DNA having a 3' terminal sequence, d. selectively hybridizing a second primer with a portion of the 3' terminal sequence of the single-stranded DNA, wherein the second primer contains a functional group capable of subsequent

5 manipulation, and e. contacting the partially digested primer-DNA hybrid with a DNA polymerase to form a contiguous double-stranded DNA containing the functional group, thereby adapting the double-stranded DNA for subsequent manipulation. An example of this method is provided in Fig. 4.

10 An example of the first primer, or zipper gene specific primer (zgsp) is given in Fig. 1.A. The zgsp's 5' terminal region is preferably between 6-20 bases especially between 6-12 bases. The 3' region is preferably 20-24 bases but may be 14 to 60 bases. Partial digestion with the 5' exonuclease generally can take between 5 seconds and 1 1/2 minutes, but preferably is allowed to proceed for between 30 seconds and 1 minute. An  
15 example of the second primer, or zipper adapter primer (zap) is shown in Fig. 1.B. Preferably, the 5' sequence of the zap is between 6 and 100 bases, more preferably 10 to 40 bases. The 3' sequence may generally be between 6 to 12 bases, with 12 bases preferred. Partial digestion with the exonuclease generally occurs with the range of 5 seconds to 1 1/2 minutes, especially within the range of 30 seconds to 1 minute. The 5'  
20 exonuclease is T7 gene 6 in a preferred method. The functional group may be a detectable moiety selected from the group consisting of fluorescein, digoxigenin, biotin, radioactive isotopes, acridine-esters, and enzymes. Alternatively, the functional group can be selected from the group consisting of GC clamps, RNA transcription promoters, M13 sequencing primer sites, the Kozak consensus sequence for protein expression,  
25 biotin, uridines, and DNA binding protein recognition sites. Also, the functional group may be DNA encoding a restriction endonuclease recognition site.

It is also a purpose of this invention to provide a method of adapting a nucleic acid for sequencing comprising: a. amplifying the nucleic acid using a first primer  
30 containing deoxyuridine monophosphate and a second primer not containing deoxyuridine monophosphate to yield a double-stranded DNA with a polyuridylylated 5'

end on one strand, b. partially disassociating the resultant double-stranded DNA with uracil DNA glycosylase to form a 3' overhang on the polyuridylated 5' end, c. digesting the double-stranded DNA with an 3' exonuclease which digests only the 3' strand from the end opposite the polyuridylated 5' end for a time sufficient to convert at least a  
5 portion of the double-stranded DNA to a single-stranded DNA, d. extending the partially digested 3' strand with a polymerase such that sequencing by a dideoxynucleotide process can be performed on the single-stranded DNA. An example of this method is shown in Figure 14.

10 The preferred 3' exonuclease is exonuclease III. The preferred number of uracils to be attached to the polyuridylated 5' end is between 4 and 6 but should be at least 3. A particularly useful exonuclease is exonuclease III. The time recommended for digestion is between 1 and 10 minutes, but may be between 30 seconds and 30 minutes. The invention also provides that digestion may continue for a time sufficient to convert the  
15 entire double-stranded DNA to single-stranded DNA, followed by binding the single-stranded DNA with a nucleic acid primer which selectively hybridizes with the single-stranded DNA, such that sequencing by a dideoxynucleotide process can be performed on the single-stranded DNA. Depending upon the length of the DNA and the location of the primer hybridization region, this time will usually be greater than 10 minutes.  
20 This alternative method is shown in Figure 15.

Furthermore, the invention provides a method of adapting a double-stranded DNA for insertion into a vector comprising digesting the double-stranded DNA with a non-polymerasing DNA exonuclease for a time sufficient to create cohesive single-  
25 stranded DNA terminal regions which will hybridize with a vector having complementary cohesive single-stranded regions. The amount of time required is between about 5 seconds and 90 seconds. In this method, the non-polymerasing DNA exonuclease is preferably T7 gene 6. This method allows for digesting the double stranded DNA in a solution that has not been purified by the removal of free  
30 deoxynucleotides, since the exonuclease is non-polymerasing.

The present invention also provides a method of inserting a target DNA into a target vector for cloning comprising: a. annealing to the target DNA a first primer having a 3' terminal region homologous to a portion of the target DNA, and a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the denatured DNA a second primer having a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, and having a 5' terminal sequence homologous to single-stranded regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the cohesive single-stranded insert DNA regions with the target vector having homologous single-stranded regions to create a circular gapped vector-insert DNA hybrid, and h. transfecting the circular gapped vector-insert DNA hybrid into a host cell. This overall method is diagrammed in Figures 7 and 8.

An example of the first primer, or zipper gene specific primer (zgsp), is given in Fig. 1.A. The zgsp's 5' terminal region is preferably 6-12 bases. The 3' region is preferably 20-24 bases but generally may be 14 to 60 bases. Partial digestion with the 5' exonuclease generally may take between 5 seconds and 1 1/2 minutes, but preferably is allowed to proceed for between 30 seconds and 1 minute. An example of the second primer, or zipper adapter primer (zap) is shown in Fig. 1B. Preferably, the 5' sequence of the zap is between 6 and 100 bases, more preferably 10 to 40 bases, and the 3' sequence generally may be between 6 to 12 bases, with 12 bases preferred. Partial digestion with the exonuclease generally occurs with the range of 5 seconds to 1 1/2 minutes, especially within the range of 30 seconds to 1 minute. The method can be practiced by transfecting the DNA hybrid to a host cell for repairing nucleotide gaps by naturally occurring enzymes.

Furthermore, the invention provides a method of adapting a target vector for subsequent recombination comprising: a. annealing to a target DNA a first primer having (1) a 3' terminal region homologous to a portion of the target DNA (2) a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization and (3) a sequence encoding a restriction endonuclease site located at the junction of the 3' terminal and the 5' terminal region of the primer, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the denatured DNA a second primer having a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, and having a 5' terminal sequence homologous to single-stranded regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the cohesive single-stranded insert DNA regions with the target vector having homologous single-stranded regions to yield a circular plasmid, h. transfecting the plasmid into a host cell, i. culturing the host cell containing the plasmid, j. purifying the plasmid from a lysate of host cell culture, k. contacting the plasmid with a restriction endonuclease which cleaves the site originating from the first primer such that the DNA originating from the first primer remains on the target vector, and l. exposing the target vector to an exonuclease for a time sufficient to create cohesive single-stranded DNA regions originating from the first primer on the vector, thereby adapting the vector for subsequent recombination. This method is generally shown in Figure 10.

25

An example of the first primer, or zipper gene specific primer (zgsp), is given in Fig. 1.A. The zgsp's 5' terminal region is preferably 6-12 bases and the 3' region is preferably 20-24 bases but may be 14 to 60 bases. Partial digestion with the 5' exonuclease may generally take between 5 seconds and 1 1/2 minutes, but preferably is allowed to proceed for between 30 seconds and 1 minute. An example of the second primer, or zipper adapter primer (zap) is shown in Fig. 1.B. Preferably, the 5' sequence

30

of the zap is between 6 and 100 bases, more preferably 10 to 40 bases, and the 3' sequence may be between 6 to 12 bases, with 12 bases preferred. Partial digestion with the exonuclease generally occurs with the range of 5 seconds to 1 1/2 minutes, especially within the range of 30 seconds to 1 minute. In one embodiment of this method the restriction endonuclease is Apa 1. The invention contemplates that DNA repair generally occurs through native repair mechanisms.

Finally, the invention provides a method of adapting an insert DNA for subsequent recombination with a vector comprising: a. annealing to a target DNA a first primer having a 3' terminal region homologous to a portion of the target DNA, and a 5' terminal region not homologous to the target DNA and of a length sufficient for future selective hybridization, b. amplifying the annealed target DNA and first primer to yield double-stranded DNA, c. denaturing the double-stranded DNA, d. hybridizing to the denatured DNA a second primer having (1) a 3' terminal sequence homologous to the 5' terminal region of the amplified DNA, (2) a 5' terminal sequence homologous to single-stranded regions of the target vector and not homologous to the target DNA or the 5' terminal region of the amplified DNA, and (3) a sequence encoding a restriction endonuclease site located at the junction of the 3' terminal sequence and the 5' terminal sequence of the second primer, e. amplifying the hybridized sequences to yield double-stranded insert DNA, f. digesting the double-stranded insert DNA with an exonuclease for a time sufficient to convert those sequences originating from the second primers to cohesive single-stranded insert DNA regions, g. hybridizing the cohesive single-stranded insert DNA regions with the target vector having homologous single-stranded regions to yield a plasmid, h. transfecting the plasmid into a host cell, i. culturing the host cell containing the plasmid, j. purifying the plasmid from a lysate of host cell culture, k. contacting the plasmid with a restriction endonuclease which cleaves the site originating from the first primer such that the DNA originating from the first primer remains on the insert DNA, and l. exposing the target vector to an exonuclease for a time sufficient to create cohesive single-stranded DNA regions originating from the first primer on the vector, thereby adapting the insert DNA for subsequent recombination with a vector. This method is depicted in Figure 11.

In one embodiment of this method the restriction endonuclease is Pvu 1. An example of the first primer, or zipper gene specific primer (zgsp), is given in Fig. 1.A. The zgsp's 5' terminal region is preferably 6-12 bases and the 3' region is preferably 20-24 bases but may generally be 14 to 60 bases. Partial digestion with the 5' exonuclease  
5 may generally take between 5 seconds and 1 1/2 minutes, but preferably is allowed to proceed for between 30 seconds and 1 minute. An example of the second primer, or zipper adapter primer (zap) is shown in Fig. 1.B. Preferably, the 5' sequence of the zap is between 6 and 100 bases, preferably 10 to 40 bases, and the 3' sequence may generally be between 6 to 12 bases, with 12 bases preferred. Partial digestion with the  
10 exonuclease generally occurs with the range of 5 seconds to 1 1/2 minutes, especially within the range of 30 seconds to 1 minute.

Briefly stated, the present invention provides methods to improve analysis of amplified nucleic acids. Three types of critical DNA analytic procedures are improved:

15

1. Specific detection of DNA sequences by colorimetric hybridization
2. Cloning of amplified DNA into vectors without the use of restriction enzymes or DNA ligase enzyme
3. Sequencing of amplified DNA

20

In each case, partial or full digestion of double-strand PCR product with specific exonuclease under specific conditions is found to enhance analytic procedures. It is  
25 found that adapters can be usefully combined with exonuclease-treated DNA to improve capture of DNA molecules to solid support, and to enhance recombination of nucleic acids. Such exonuclease and adapter-mediated recombination of molecules can be used to incorporate diverse functional groups into DNA, including but not limited to useful nucleotide sequences (40-base GC clamps, 17-base RNA promoters, restriction  
30 endonuclease recognition sites, M13 sequencing primer sites, the Kozak consensus sequence for protein expression, etc.), chemically modified nucleotides (containing

biotin, fluorescein, digoxigenin, uradines, etc.), and DNA binding protein recognition sites (such as that for GCN4).

In DNA diagnostics applications, this invention converts double-stranded DNA to single strands by incubation with exonuclease. T7 gene 6 (an example of a 5' exonuclease) is a preferred enzyme for this application, but exonucleases with either 5' or 3' activity that selectively bind to double-strand DNA, and digest the DNA to single strands without digesting the resulting single-strand DNA, can be used. The resulting single-strand DNA is then highly reactive with oligonucleotide probes, which can then be used to capture DNA to solid support (microtiter plates, coated magnetic or agarose particles, treated nylon or glass, etc.). Each step proceeds efficiently and rapidly in standard buffers used for PCR. Following conversion to single strands and specific capture, a second oligonucleotide can be used to detect the complex. In principle, the hybridization assay is similar to other colorimetric hybridizations previously used to detect denatured double-strand DNA. However, simple conversion to single strands by exonuclease is found to enhance colorimetric hybridization signal by 100-fold (Example 1), resulting in greater specificity, and is more convenient and automatable than procedures commonly used to denature PCR products.

For DNA cloning, the current invention provides a convenient method to introduce DNA into plasmid and bacteriophage vectors without using ligase, is independent of T4 DNA polymerase, does not require construction of special vectors, and does not require extensive purification of PCR product. The method also provides a means of joining DNA to adapter oligonucleotides bearing functional groups such as biotin, fluorescein, GC clamps, RNA promoters, etc. In this method, PCR primers are synthesized to which two different defined sequences (named "zippers," 6-12 bases in length) (Fig. 1.A.) can be synthesized at the 5' of gene-specific primer pairs (one defined zipper sequence at the 5' of all upstream gene-specific primers, the second at the 5' of all downstream gene-specific primers). After amplification, the product is treated briefly with 5' exonuclease (such as T7 gene 6) to create cohesive 3' overhangs. Adapter oligonucleotides of variable length, ranging from 6 to 100 nucleotides, can be



annealed to the defined zipper sequence, and then extended by polymerase (e.g., Klenow fragment, T7 DNA polymerase, Taq polymerase, etc.). In this way, common sequences can be catalogued and then conveniently affixed to PCR products according to individual needs of particular experiments. Sets of adapters can be constructed  
5 containing sequence matching vector insertion sites; after affixing vector sequences to PCR fragment via exonuclease-extension, PCR fragment is easily cloned to the desired vector by (1) homologous recombination or (2) further treatment with exonuclease. Step 2 can be accomplished by 5' or 3' exonuclease in presence or absence of dNTPs. Individual PCR fragments can be shuttled into heterogeneous vectors by affixing  
10 different sets of adapters. Once cloned, inserts can be released from vector, affixed to a new set of adapter, and recloned into a different vector without reamplification. Alternatively, the same reaction may be used to affix adapters to vector in preparation to accept zipper-containing PCR products. The same reaction used to recombine insert with vector can join several different DNA inserts together, or affix sequences and  
15 modified nucleotide groups to inserts in vitro.

The present invention also provides improved methods for sequencing DNA. In one embodiment, a highly reliable 3' overhang is produced at one end of a DNA molecule by attachment of an oligonucleotide containing uridines, then treating with  
20 uridine DNA glycosylase (UDG). Immediately following, the DNA molecule is digested from the opposite 3' end with exonuclease III for variable lengths of time. After stopping the reaction and removing of exonuclease, dideoxy sequencing can commence from the variably digested end. This application of the invention results in greatly extended ability to sequence DNA ends, permitting >800 bases of sequence to  
25 be obtained per reacted DNA.

Thus, the invention offers combinations of exonuclease enzymes, zipper-defined sequences on PCR primers, and adapter oligonucleotides to produce a highly flexible, unified approach for cloning, sequencing, expressing, and detecting DNA. The catalog  
30 of primers created for this integrated system can be used with extreme economy. These

and other aspects of the present invention will become evident upon reference to the following detailed description and attached drawings.

The following examples are offered by way of illustration, not by way of  
5 limitation.

**Example 1.** *Detection of parvovirus B19 by the exonuclease amplification confirmation capture technique (EXACCT).*

10 The strategy shown (Fig. 2) was used to detect parvovirus B10, the etiologic agent of *Erythema infectiosum*<sup>25-26</sup> from a bank of tissue or serum specimens using primers and amplification procedure previously described<sup>27,28</sup>, yielding a 284 bp product. Following PCR amplification, three identical reaction aliquots (320 ml or ~400 mg DNA) were digested with exonuclease for 15 min at room temperature,  
15 electrophoresed, and transferred to nylon under nondenaturing conditions. To test the efficiency of exonuclease digestion-dilution, undigested PCR product was compared to product digested with 4 U/ml exonuclease, 0.4 U/ml, and 0.04 U/ml, and to single-strand PCR product produced using reduced P1 and excess P6 in asymmetric PCR. After gel electrophoresis, these DNA products were assessed by ethidium staining and  
20 by nondenaturing Southern blot analysis. The agarose gel used to size separate DNA was not exposed to alkali prior to transfer, Southern blot capillary, so that only single-strand DNA would be efficiently detected by the digoxigenin-labeled oligonucleotide probe. Double-strand DNA not exonuclease-treated was poorly detected compared to product treated by exonuclease or amplified by asymmetric PCR.

25

These same DNAs were detected (Fig. 6) in streptavidin-coated plates. Twenty ml of PCR product was added to 200 ml of hybridization buffer (4xSSC, 20 mM Hepes, 2 mM EDTA, 0.15% Tween20) containing both biotinylated (100 ng/ml) and digoxigenin-labeled (200 ng/ml) probes. Hybridization reactions were performed at  
30 45°C for 1 h. As in other colorimetric detection protocols previously in the art, double-stranded PCR products were heat-denatured at 95°C for 5 min prior to hybridization.

Single-stranded PCR products produced by asymmetric PCR or by exonuclease digestion were not heat-denatured. Equal amounts of control DNA amplified from an unrelated region of the B19 gene were treated with exonuclease in parallel. Following hybridization, 100 µl of each mixture was added to duplicate from an unrelated gene  
5 that was exonuclease-treated in parallel with experimental samples. Following hybridization, 100 µl of each mixture was added to duplicate wells of a streptavidin-coated microtiter plate prepared by coating with 200 µg of biotinylated BSA in 100 µl PBS/well overnight at 4°C, washed with PBS containing 0.15% Tween20, then saturated with 1000 µg of streptavidin in 100 µl PBS with 0.5% gelatin/well for 1 h at  
10 room temperature with shaking. The plate was then washed and 100 µl/well of 3,3',5,5'-tetramethyl-benzidine chromogen solution was added and incubated at room temperature. The color development was stopped after 15 min with 100 µl/well 2 M phosphoric acid, and the plates read at 450 nm in a MR5000 Microplate reader (Dynatek, Chantilly, VA).

15

Results of this experiment showed that single-strand PCR product produced by exonuclease treatment was detected in this format with 100x sensitivity of the denatured double-strand PCR product. Signal (in OD units) for DNA treated with exonuclease (T7 gene 6, .04 U/ml - 4 U/ml) varied from 1.3 to >2.0 OD units. Signal for denatured  
20 double-strand DNA was ~.15 OD units. Control DNA did not give appreciable signal, whether heat- or exonuclease-treated.

After establishing that the procedure enhances PCR signal using a wide range of exonuclease activities, we tested the invention on a panel of geographically dispersed  
25 serum samples collected from seronegative and seropositive individuals suspected of being infected by parvovirus B19 on clinical grounds. This pool was also studied by single-round PCR (35 cycles total) followed by Southern blot analysis, and by extended PCR with nested primers (60 cycles total). B19 copy number in the infected sera (20 cycles total) was estimated by comparison to parallel amplification of known amounts  
30 of B19 DNA diluted in serum. Samples positive by single-round PCR were found to contain 350-3,500 B19 genomes, while those negative by single PCR but positive by

nested PCR contained as few as 3-30 copies of B10. In side-by-side comparison, we have found that the invention detected 3-30 copies of B10 after a single PCR from these clinical specimens, and from known amounts of B19 DNA diluted in serum. Single round PCR followed by the invention was at least 100-fold more sensitive than single-  
5 round PCR alone, 10-50 times as sensitive as Southern analysis of single-round PCR products, and equivalent in sensitivity to nested PCR. The invention required considerably less time, effort, and technical skill than Southern blot analysis of PCR products (<3 h for the invention versus >2 days for Southern analysis) and was more specific, since two separate probes are required to hybridize to target DNA in EXACCT  
10 detection. Compared to nested PCR, the risk of contamination was dramatically lessened, since PCR tubes are not opened between PCR rounds. The risk of false-positive carryovers due to carryover of PCR-amplified DNA can be further reduced by UNG treatment, whereas UNG cannot be used in both rounds of a nested PCR. Signal detection also appeared to correlate more quantitatively with the number of B19  
15 genomes initially present in samples compared to nested PCR, which gave the same maximal signal regardless of the amount of B19 present.

The results demonstrate greatly enhanced ability to make sensitive, specific, and convenient diagnosis of human parvovirus B19 infection in seronegative  
20 immunocompromised individuals. DNA detection is done simply without gel electrophoresis on blot analysis in microtiter plates, a familiar, convenient format for clinical laboratories that can be easily automated. This procedure can easily detect a large number of other viruses and microorganisms by PCR, differentiate PCR-amplified polymorphic human DNA, and detect mutations in genetic disease.

25 Modifications of this strategy (Fig. 3) have been useful in protecting selected DNA strands, such that one strand is digested completely and the other is left intact. A tail of 5' uridines was synthesized at the 5' of the downstream parvovirus B19 amplifying primer. After PCR, the resulting DNA (300 mg) was treated in the same  
30 reaction vial used for amplification (without changing buffer) with *E. coli* UDG (10 U/reaction) (Bethesda Research Laboratories, Bethesda, MD), to create a 3' overhang.

The uracil containing end of the DNA was thus protected against 3' exonuclease attack during digestion with ExoIII (10 U/reaction, 37°C for 10 min) (New England Biolabs, Wilmington, DE). The resulting single-strand DNA was then hybridized under the same conditions described for format A above (Fig. 2), and signal detected in the microtiter plate. Results were identical to detection with T7 gene 6, demonstrating that both 3' and 5' exonucleases can successfully enhance DNA detection by hybridization. In Example 1, digestion of both strands by a single enzyme in format A produced the desired effect. For detection of shorter DNAs, it is evident that format B might be advantageous to protect the full length of one strand, such that sufficient length of DNA is presented to hybridizing capture and detection probes.

**Example 2.** *Detection and speciation of mycobacteria based on exonuclease-coupled amplification.*

The method of Example 1 was applied to detect mycobacteria genes amplified from clinical specimens. Two different mycobacterial genes were targeted, including *Mycobacterium tuberculosis* (MTB) IS6110 transposon and the gene for the mycobacterial 16S ribosomal (r)RNA subunit. Sample preparation, amplification procedures, and PCR primers were as previously described for the MTB IS6110 by Eisenach et al.<sup>30</sup>. The 16S ribosomal amplifications were as previously described<sup>31</sup> based on sequence catalogued<sup>32,33</sup>. After amplification, PCR products (20 µl, containing ~50 ng DNA) were treated with T7 gene 6 exonuclease (.4 U/sample) then hybridized to capture and detect oligonucleotides using conditions described in Example 1. Sequence of capture probe for the IS6110 was digoxigenin-labeled probe 5'-Dig-AGTAATTCCGGACAACGCTC (SEQ ID NO:3) and biotin-labeled 5'-Bio-TTTCACGAACAACGCGACAA (SEQ ID NO:4). Sequence of capture probe for 16S rRNA PCR products was digoxigenin-labeled 5'-Dig-ACGGTGCCCGCAAAGTGTGG (SEQ ID NO:5) and biotinylated 5'-Bio-ACCGTGAGGGCATCGAGGTG (SEQ ID NO:6). Results demonstrated the extreme sensitivity of this procedure in specifically identifying PCR-amplified DNA. Control DNA amplified from unrelated genes were not detected by this procedure.

The flexibility of the system has been tested using different materials as solid support matrix. The exonuclease-treated DNA oligonucleotide hybrids have been captured to streptavidin-coated magnetic microspheres (DynaM-280 particles, Oslo, Norway) and to streptavidin-coated microspheres packed in columns and pipette tips (U.S. Biochemical, Cleveland, OH). In each case, signal detection is markedly enhanced by the step of converting the DNA into single strands via exonuclease treatment. Other streptavidin-coated support matrices are currently under investigation, including streptavidin-coated glass, streptavidin-coated microtiter prongs, and streptavidin-coated nylon. It is thought that the mechanism by which the capture oligonucleotide is bound to solid support is not important, and that oligonucleotides covalently linked to support matrices without streptavidin should be equally effective. We have also demonstrated that oligonucleotides can be bound to matrix prior to hybridization of digested DNA.

**Example 3. Detection of mycobacterial rRNA without PCR.**

Nucleic acids expressed at high copy numbers can be detected in cells without in vitro amplification. In this example, mycobacteria culture grown on agar slant or in BACTEC<sup>34</sup> broth were lysed by sonication using the procedure of Drake and others<sup>35,36</sup>. Immediately following the capture and detection probes described in Example 2 for 16S rRNA were hybridized to the released mycobacterial 16S rRNA. Since it is known that such rRNA is present at 15,000 copies per bacterium, it was felt that amplification would not be necessary to detect this nucleic acid<sup>32</sup>. For this example, the detection oligonucleotide was <sup>32</sup>P-labeled by terminal deoxytransferase using standard procedures. 16S rRNA oligonucleotide hybrids were captured to streptavidin-coated microspheres. Results showed that the capture-detection technique clearly distinguished mycobacteria from other bacteria such as pseudomonas (a frequent contaminant of mycobacterial cultures) from both broth and culture. Other techniques, such as Gen-Probes (Gen-Probe, San Diego, CA) also can detect mycobacteria grown on agar slants via hybridization to species-specific acridine probes. However, substances from broth and in clinical specimens are found to interfere with detection by

Gen-Probes; thus, the Gen-Probes system has not proved useful in early mycobacteria detection<sup>35</sup>. The current invention overcame this limitation by capturing the single strand 16S rRNA to solid support, so that inhibitors are washed away prior to detection. The capture of nucleic acids to solid support by the present invention can enhance  
5 detection by numerous different kinds of probes, including acridine-labeled probes.

The potential of detecting other high copy number nucleic acids by the invention without amplification can be performed. Mitochondrial DNA can be linearized by restriction endonuclease digestion or by sonication. Then, linearized double-strand  
10 DNA can be converted to single strands via exonuclease, captured to solid support, and detected with second labeled oligonucleotide. Some species of mRNA or genomic DNA could also be detected by this invention without amplification.

**Example 4.** *Recombination of human G protein cDNA to several different adapters  
15 bearing functional groups and vector-complementary sequence.*

Figure 4 depicts a General Ligation Using Exonuclease-Extension (GLUEE). This procedure has been used to join more than 15 different cDNAs coding for segments of human guanine nucleotide binding ("G") proteins and G protein coupled  
20 receptors to several different adapter oligonucleotides bearing functional sequences and modified chemical groups. In each case, the cDNA to be recombined was amplified using primers consisting of 20 bases specific to target (zipper-containing gene-specific sequence, or zgsp), Fig. 1.A., and continuing at the 5', an additional 6-12 bases referred to as "zippers." Each upstream gsp PCR primer was synthesized with a portion of  
25 "zipper forward" at the 5' (CGAGGGAAGAGG) (SEQ ID NO:1)= the underlined portion representing the minimal sequence to be attached). Each downstream PCR primer was synthesized with a portion of "zipper reverse" (CGCACGCGGGAG) (SEQ ID NO:2). After amplification, each PCR fragment had incorporated these zipper sequences at the 5' that could be used as targets for further amplification or to create  
30 cohesive ends. "Zipper adapters" (ZAPs, Fig. 1.B.) refer to variable-length oligonucleotides containing the zipper sequences at the 3' and commonly used

sequences and chemical groups following. Thus, the 3' of the ZAPs overlap matching zipper sequences present at the 5' of the gene-specific PCR primers. The zipper containing amplified DNA is treated briefly with a 5' exonuclease (such as T7 gene 6). After generating cohesive ends of variable length, ZAPs can be hybridized. At room temperature, we have found that 6-base-complementarity between cohesive-end DNA and ZAP primers are sufficient to assure good hybridization. After hybridization, the adapter primer can initiate fill-in DNA synthesis via fresh DNA polymerase. DNA is extended in both directions, leading to double-strand DNA including the sequences and chemically modified nucleotides uncompounded into the ZAP.

10

The aspect of the invention described in this example has been used on numerous DNAs under varying conditions. PCR primers can be removed prior to exonuclease extension via microcon concentration or via ethanol precipitation. Initially, this step was designed so that remaining zgsp primers do not compete with ZAPs in exonuclease extension. However, in most cases we have found removal of PCR primer unnecessary. Untreated amplification product can be digested with T7 gene 6 exonuclease (.4 U/reaction) in standard PCR buffers for 1 min at room temperature<sup>23,24</sup>. After this digestion, the exonuclease is heat-inactivated at 70°C for 5 min. ZAP primers are then added (.5-5 pmole) and annealed to DNA for 10 min at room temperature.

20 Fresh DNA polymerase (Klenow fragment, Taq polymerase, etc.) is added, again without changing buffers.

This invention was used to join adapters to 400, 700, and 1,100 bp segments of human G<sub>q</sub> and G<sub>11</sub> protein cDNA amplified from cell lines derived from small cell lung tumors and from human umbilical cord endothelial (HUVEC) cells. Several different adapters were used (Table 1). One set of adapters (FAQ and RAQ) were used to introduce sequence complementary to the multicloning insertion site of protein expression vector QE9. Another set of adapters (FAB and RAB) introduce sequence complementary to the multicloning site of pET protein expression vectors. These adapters proved useful in several different ligase-free cloning strategies, ultimately

30



permitting expression of amplified cDNA with 6-histidine amino terminal tags that can be used in protein affinity purification strategies.

- One set of adapters (FAQ and RAQ) were used to introduce sequence
- 5 complementary to the multicloning insertion site of protein expression vector QE9. Another set of adapters (FAB and RAB) introduce sequence complementary to the multicloning site of pET protein expression vectors. These adapters proved useful in several different ligase-free cloning strategies, ultimately permitting expression of amplified cDNA with 6-histidine amino terminal tags that can be used in protein
- 10 affinity purification strategies.

**Table 1. Sequences of oligonucleotides**Zipper sequences*zipper* Forward CGAGGGAAGAGG (SEQ ID NO:1)*zipper* Reverse CGCACGCGGGAG (SEQ ID NO:2)

5

Zipper adaptersGC-For CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCGGGCCC  
GAGGGAAGAGG (SEQ ID NO:7)10 RF-F AGGGAGACCGGAATTCGGATCCCATATGGCGGCCGCGATCGAGG  
GAAGAGG (SEQ ID NO:8)RE-R GACAGAGCACA GAATTCGGATCCAAGCTTCGATCGCACGCGGGAG  
(SEQ ID NO:9)U-FOR CAUCAUCAUCAUCAUGCGATCGAGGGAAGAGG (SEQ ID  
NO:10)

15 U-REV CUACUACUACUACGAT CGCACGCGGGAG (SEQ ID NO:11)

FAB CATCATCACAGCAGCGGCCTGGTGCCGCGCGGCGGCGGATCGAG  
GGAAGAGG (SEQ ID NO:12)RAB TCAGCTTCCT TTCGGGCTTTGTTAGCAGCCGGATCCGATCGCAC  
GCGGGAG (SEQ ID NO:13)20 FAQ-9 AACTATGAGAGGATCGCATCACCATCACCATCACGGATCCTCGATC  
GAGGGAAGAGG (SEQ ID NO:14)RAQ-9 GGATCTATCAACAGGAGTCCAAGCTCAGCTAATTAAGCTTCG  
ATCGCACGCGGGAG (SEQ ID NO:15)Zap-F25 TGTAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAA  
TTCGATCGAGGGAAGA GG (SEQ ID NO:16)Zap-R AACAGCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTA  
TAGAATACTCAAGCTTGGATCCGATCGCACGCGGGAG (SEQ ID  
NO:17)

30

Examples of Gene specific primers (gsp)

AT-F cgagggaagaggTACAGCATCATCTTTGTGGTGGGGA (SEQ ID NO:18)

AT-R cgcacgcgggagTCGAATTCGAGACTCATAATGA (SEQ ID NO:19)

Ren-F cgagggaagaggTAGGTCAGCAACATGGACTATGT (SEQ ID NO:20)

35 Ren-R cgcacgcgggagTAGCGGGCCAAGGCGAACC (SEQ ID NO:21)

ACE-F cgagggaagaggTGCCTCCCCAACAAAGACTGCCA (SEQ ID NO:22)

ACE-R cgcacgcgggagTCCACATGTCTCCAGCAGATG (SEQ ID NO:23)

GAP-F cgagggaagaggTCCATGGAGAAGGCTGGGG (SEQ ID NO:24)

GAP-R cgcacgcgggagTCAAAGTTGTCATGGATGACC (SEQ ID NO:25)

40 TMB-F cgagggaagaggTCCNYTNYTNYTNGGNGCNGCNATGGC (SEQ ID NO:26)

TMB-R cgcacgcgggagTANACCCANGGRTCNARDATYTGRTTCCA (SEQ ID  
NO:27)

BmTM3-F cgcccgccgcgCCCCGCGCCC (SEQ ID NO:28)

BmTM7-R cgcacgcgggagTANARNGCRAANGGRTTNACRCA (SEQ ID NO:29)

45 XREN-F cgagggaagaggTTAYGGNGARATHGGNATHGGNACNCC (SEQ ID  
NO:30)

- XREN-R cgacgcgggagTYTGNGTNACNGT DATNCCNCCNACNGT (SEQ ID NO:31)  
 XGA2-F cgagggaagaggTGCNGGNGARWSNNGGNAARWSNAC (SEQ ID NO:32)  
 XGC2-r cgacgcgggagTCNSWNCKYTGNCCNACNACRTC (SEQ ID NO:33)  
 5 Go-f cgagggaagaggATGGGATGTACTCTGAGCGCAGAGG (SEQ ID NO:34)  
 Go-r cgacgcgggagCGCACGCGGGAGTCAAAGTTGTCATGGATGACC (SEQ ID NO:35)  
 Gq-F cgagggaagaggATGACTCTGGAGTCCATCATGGCGT (SEQ ID NO:36)  
 hGs cgacgcgggagTCCATGGAGAAGGCTGGGGcc (SEQ ID NO:37)  
 10 mG15 cgagggaagaggATGGCCGGCTGCTGCTGTTTGTCTG (SEQ ID NO:38)  
 mG11 cgagggaagaggATGACTCTGGAGTCCATGATGGCGT (SEQ ID NO:39)

- \* Upper case denotes sequences matching PCR template targets. Lower case denotes additional sequence added to facilitate zipper PCR strategies. Sequence is given in IUPAC nomenclature; in some cases, more than one nucleotide is synthesized to permit amplification of multiple gene sequences belonging to a gene family.

#### Vector adapters

- cZR-RP1 CTCCCGCGTGC GATCGG TCATAGCTGT TTCCTGTG (SEQ ID NO:40)  
 20 ZR-ada CACAGGAAAC AGCTATGACC GATCGCACGC GGGAG (SEQ ID NO:41)  
 cpet ir CTCCCGCGTG CGATC GGGCTGCTGCCACCGCTGAGC (SEQ ID NO:42)  
 ZR-adet GCTCAGCGGT GGCAGCAGCC CGATCGCACG CGGGAG (SEQ ID NO:43)  
 25 pET-y CGGCACCGTC ACCCTGGATGCTGTAGG (SEQ ID NO:44)

#### Vector amplifying primers

- M13-FZ CGA GGG AAG AGG TGGCGATTAAGTTGGGTAACGCC (SEQ ID NO:45)  
 30 M13-RZ GGC ACG CGG GAG TCACACAGGAAACAGCTATGAC (SEQ ID NO:46)  
 M13-RB(ZF) CGA GGG AAG AGG TGTGTGGAATTGTGAGCGG (SEQ ID NO:47)  
 35 -20 forward GTAAAACGACGGCCAGT (SEQ ID NO:48)  
 -47 forward CGCCAGGGTTTTCCCAAGTCACGAC (SEQ ID NO:49)  
 reverse-48 GAGCGGATAACAATTTACACAGG (SEQ ID NO:50)  
 M13-FZ CGA GGG AAG AGG TGGCGATTAAGTTGGGTAACGCC (SEQ ID NO:51)  
 40 M13-RZ GGC ACG CGG GAG TCACACAGGAAACAGCTATGAC (SEQ ID NO:52)  
 M13-RB(ZF) CGA GGG AAG AGG TGTGTGGAATTGTGAGCGG (SEQ ID NO:53)  
 pZO4 upper CGAGGGAAGAGGGGCCccccccccgggttaaaaaaaaaaaaaa  
 45 (SEQ ID NO:54)

pZO4 lower CGACGTTTTT TTTTTTTAAA CCCGGGGGGG GGGGGGCC (SEQ  
ID NO:55)  
f/r join CGA GGG AAG AGG T GGC ACG CGG GAG T (SEQ ID NO:56)  
GT<sub>10-F</sub> TGT AAA ACG ACG GCC AGT AGC AAG TTC AGC CTG GTT AA  
5 (SEQ ID NO:57)  
GT<sub>10-R</sub> TCC CAG TCA CGA CGT TGT GGG GTA AAT AAC AGA GGT GGC  
(SEQ ID NO:58)  
GT<sub>11-F</sub> GGT GGC GAC GAC TCC TGG AGC CCG (SEQ ID NO:59)  
GT<sub>11-R</sub> TTG ACA CCA GAC CAA CTG GTA ATG (SEQ ID NO:60)  
10 X<sub>t3</sub>: CGA CGG CCA GTC GAC TCT AGT TTT TTT TTT TT (SEQ ID  
NO:61)  
X<sub>t4</sub>: CGA CGG CCA GTC GAC TC (SEQ ID NO:62)

The most flexible adapter sets that have proven most useful in multiple  
15 applications are ZAP1-F and ZAP1-R (SEQ ID NOS:16 AND 17) (Fig. 5). ZAP1-F  
attaches to cohesive ends produced by 5' exonuclease digestion of DNA containing the  
zipper forward sequence. After extension is complete, an M13 universal sequencing  
site, a T7 RNA promoter and several restriction sites are incorporated to the DNA.  
ZAP1-R attaches to the zipper reverse sequence to incorporate an sp<sub>6</sub> RNA promoter, an  
20 M13 reverse sequencing site, and additional restriction endonuclease sites. ZAP-P\* and  
ZAP-R\* represent adapters with the same sequence into which biotin has been  
incorporated at the 5'. By joining a biotinylated oligonucleotide to one end of DNA and  
an unbiotinylated oligonucleotide to the other, extended DNA can be bound to  
streptavidin-coated solid support (Fig. 6). Several different manipulations can then be  
25 performed in the solid phase, including cRNA transcription of either strand, and  
sequencing of DNA. Bead-bound DNA can be gathered, purified, and reused; multiple  
step molecular manipulations are facilitated since buffers and by-products are quickly  
exchanged during washing of other bead-bound material. DNA can be reamplified from  
its solid support or cleaved by restriction endonuclease off the bead. The invention thus  
30 extends solid phase sequencing strands of Uhlén and coworkers<sup>22</sup>. Unlike methods in  
the art, however, only two biotinylated primers are used to bind innumerable different  
DNAs to solid phase in different experiments. Since as little as 0.5 pmole of adapter  
has been used to bind sufficient DNA to solid phase for these applications, a 1 pmole  
synthesis of adapter oligonucleotide is theoretically sufficient for use in 10<sup>6</sup>  
35 experiments. Production of multiple biotinylated primers is a major expense of current

solid-phase DNA technology, especially since investigators often remake identical pairs of primers with and without biotin to bind opposite DNA strands to solid support.

In addition to supporting multiple molecular procedures, the sequence of ZAP1-F and ZAP1-R is identical to sequence present at the multiple cloning site of pGEM. Eighteen bases in each ZAP are identical to sequence in M13 and pUC vectors. These sequence identities are the basis of cloning methods discussed in Example 6.

Thus, after extending with ZAP1-F and ZAP1-R adapters, DNA can be easily purified, converted to single-strand probe, transcribed into cRNA, restriction-digested or easily cloned into pGEM-32, M13, pUC plasmid, and related vectors. These applications are all routinely used in our laboratory; over 15 different G protein and G-coupled receptor cDNAs have been manipulated by these methods.

**Example 5. *Joining of zipper-PCR products to adapters through secondary PCR.***

Figure 7 depicts the joining of DNA to ZAP primers by a second method, namely through secondary PCR (also referred to as a "piggyback PCR"). Zipper sequences at the extreme ends of PCR products become the target of adapter-primed additional PCR. Several PCR formats have been developed to accomplish this; again, cDNA from many different G proteins have been manipulated using these procedures. In Format 1, an initial round (30 cycles) of PCR are primed by gsp primers under conditions using target-determined amplification parameters. The 5' zipper additions do not appear to impair PCR; indeed, we often observe that the length of ts- primers is more efficient in PCR than similar primers lacking the 12-base zipper additions. A 1-ml aliquot of the product of this amplification is then used as the template in a second round (30 cycles) of PCR primed by the universal primers, using a standard amplification protocol (95°C, 40 s; 53°C, 45 s; 72°C, 60 s).

A second format for adapter attachment is referred to as the 10/25 protocol. This format is designed to keep the number of total PCR cycles low in order to

minimize Taq-induced sequence errors during amplification, and to minimize the use of PCR reagents. In this format, an initial round (10 cycles) of PCR is done using a reduced amount of the ts- primers (2 pmol each) in a small volume reaction (total vol - 3 ml). Immediately following the initial round, 40 ml of a fresh 1xPCR mixture is  
5 added, consisting of standard concentrations of Perkin-Elmer PCR buffer, dNTPs, Mg, fresh Taq DNA polymerase, and 20 pmol each of primers ZAP-F and ZAP-R. PCR is continued for an additional 25-30 cycles, using the same parameters outlined for the second round of format 1.

10 Although adding adapters by GLUEE and by secondary PCR can result in essentially identical products, each has advantages for different uses. Secondary PCR requires only one enzyme (Taq polymerase) in a format that is familiar to PCR users. However, much more adapter primer is required to drive secondary PCR than is necessary for GLUEE; when larger adapters are required, such as ZAP1-F or ZAP1-R,  
15 adapter ligation by GLUEE is much more economical. Since annealing in GLUEE pairs cohesive single strands, a process which proceeds avidly at room temperature, shorter zipper sequences can be used at the 5' of gsp primers (6-12 nucleotides). In contrast, annealing of adapters to denatured double-strand PCR products has been shown to require slightly more extensive overlap; less than 12 nucleotides appear to be  
20 inefficient in supporting secondary PCR. In some experiments, GLUEE and secondary PCR can be used sequentially; after ligating adapters onto DNA targets by GLUEE, additional material can be made by secondary amplification primed by additional adapter. When long adapters are required (ZAP1-R is 66 nt and ZAP1-R is 81 nt), oligonucleotide can be conserved by ligating small amounts onto DNA using GLUEE,  
25 then reamplifying using secondary PCR primed by a shorter oligonucleotide complementary to sequence at the 5' of the adapter (such as the M13 universal forward and reverse primers incorporated into ZAP1-F and ZXAP1-R.

**Example 6.** *Demonstration of directional cloning of cDNA ligated to ZAPs into*  
30 *plasmid vectors in format A.*

After joining DNA to certain ZAPs (either by GLUEE or by secondary PCR), the DNA can be introduced into vector plasmids or lambda phage. Primer pairs ZAP1-F and ZAP1-R are used to clone DNA into M13, pUV, or pGEM plasmids. Primers RAQ and FAQ introduce DNA into pQE-9, and FAB and RAB are for cloning into pET  
5 vectors. In each case, adapters introduce 320-60 bp sequences that match corresponding sequences of vectors (Fig. 8). After brief digestion with either 3' or 5' exonuclease, cohesive ends are generated in the end of vector and insert. Although we have used ExoIII, T7 polymerase (in the absence of dNTPs), and certain lots of T4 DNA polymerase, the preferred enzyme in our laboratory at present is T7 gene 6  
10 exonuclease. DNA (vector linearized by restriction enzyme digestion and insert) was treated for 30 sec to 1.5 min with 1-4 U T7 gene 6 in standard buffers used for amplification. DNA is then heated to 70°C for 10 min to inactivate enzyme. Vector and insert (10-100 mg each) are then annealed for 10-30 min at room temperature. In some cases, the annealed vector-insert is extended by the addition of fresh polymerase  
15 (Klenow fragment, Taq DNA polymerase, etc.) to fill in gapped duplex DNA; this can reduce background from extensively digested vector that inappropriately recircularized. In most cases, however, the gapped vector insert is directly used to transfer bacterial hosts without repair by DNA extension. If gapped DNA is transfected without repair, the vector-insert complex is heated to 70°C for 10 min prior to transfection to disengage  
20 any inappropriately hybridized ends. The extent of vector digestion or the presence of gaps in vector-insert duplex makes little difference in cloning efficiency, as bacterial repair enzymes repair gapped DNA after transformation. We have found on numerous occasions that several hundred clones from 100 mg input DNA can easily be achieved. Background transformation of vector without insert is low or nonexistent, depending  
25 primarily on the efficiency of restriction enzyme digestion during the step by which vector was linearized. Background can be further lowered by using two or more restriction enzymes. The system is compatible with blue-white selection in vectors containing lacZ at the cloning site.

30 The utility of this system has been demonstrated by cloning more than 20 different PCR-amplified DNA into 5 different plasmids (pGEM-3Z, M13 MP18, pUC,

pQE-9, and pET-15). Presence of insert has been confirmed by standard plasmid analysis, including restriction digestion and sequence analysis by standard sequenase dideoxysequencing protocols. Orientation of inserts was found to be determined appropriately by the direction of adapter primer joined to DNA. DNA cloned into  
5 protein expression vectors (pQE-9 and pET-15) has been appropriately expressed as 6 histidine recombinant fusion protein, and has been purified using methods suggested by the respective vector manufacturers.

It is apparent that the invention could be used to clone DNA into lambda phage  
10 vector with equal efficiency or to any other cloning vehicle. Cloning vehicles do not need to be altered; insert DNA need only be joined to adapter-containing vector complementary sequence. It is also apparent that aliquots of one reaction of PCR-generated DNA can be joined to several different vector-specific adapters and cloned simultaneously to several different cloning vehicles with specialized uses. We have  
15 constructed four different cDNA libraries from mRNA derived from small cell lung tumors after reverse transcription (RT) PCR amplification using mixed oligonucleotide primers. It is also evident that directional cloning by the invention can be used to construct cDNA libraries without PCR.

20 **Example 7. *Demonstration of directional cloning of cDNA into modified vectors cloning in format B.***

Example 6 demonstrates cloning of ZAP-joined DNA into unmodified vectors. In a variation of this procedure, we have simply introduced the 12-base zipper sequence  
25 into the plasmid vector so that primary PCR products may be cloned without the use of ZAPs. Zipper sequences have been introduced into vectors (including pQE-9, pGEM-32, pET-15, and M13 MP18) using both standard and novel techniques. Resulting modified zipper-containing vectors have been named pZQE-9, pZGEM, pZET, and pZM13, respectively. Methods of vector modification are outlined below.

30

A. Creation of pZGEM by genetic engineering (Fig. 9).



Sequences to be introduced into vector were made on an ABI 394 oligonucleotide synthesizer (Applied Biosystems, Foster City, CA). Two overlapping oligonucleotides were made, with complementary sequence containing zipper sequences, Apa 1 restriction site, and EcoR1 and HindIII compatible cohesive ends  
5 (SEQ ID NOS:63 and 64). Oligonucleotides were desalted and used without gel purification. 5' Terminal phosphates were added to oligonucleotides with T4 polynucleotide kinase and ATP under standard conditions<sup>13</sup>. After kinase treatment, oligonucleotides were heated to 70°C for 10 min; 50 ng of each oligonucleotide were annealed together at 65°C for 20 min and cooled to room temperature. The vector pQE-  
10 9 was digested with BamHI and HindIII, and the linearized vector gel-purified. Vector and annealed oligonucleotides were combined in a 10-ml reaction mix with freshly diluted ATP and DNA ligase (BRL, Bethesda, MD) using conditions specified by the manufacturer. After ligation overnight at 16°C, the DNA was transformed into competent *E. coli* HB1-01 cells (BRL), and transformed cells were plated on LA-  
15 ampicillin plates. Seventy-five colonies were found to result. Standard preparation of plasmid DNA from 3 of these colonies using DNA purification columns (Qiagen, Foster City, CA) followed by standard sequencing protocols (Sequinase, U.S. Biochemical, Cleveland, OH) demonstrated the expected sequence.

20 *B. Introduction of zipper sequence into pGEM03Z via exonuclease cloning.*

A simpler procedure was used to introduce zipper sequence into plasmid: pQE-9 simultaneously with the protein expression cloning of the mouse G protein b<sub>3</sub> subunit. In this procedure, Apa 1 restriction sites were added at the junction of zipper and gsp  
25 sequences in the amplifying primer. After amplification of the 1076 bp b<sub>3</sub> open reading frame (ORF) from a mouse plasmid clone in pGEM-32 (generously provided by Dr. Mike Levine, Johns Hopkins University, Baltimore, MD), the cDNA was cloned into Bam-HindIII digested pQE-9 linear vector using methods outlined above. After the resulting plasmid was selected, grown, and purified, the insert was released by digesting  
30 at the introduced Apa 1 site flanking both ends of insert (Fig. 10). After gel purification, the linearized vector was demonstrated to have zipper sequence present at

each end.

C. Cloning of DNA into zipper-containing vector.

5        Zipper sequences introduced into vectors by conventional or exonuclease  
cloning methods are available for cloning. Procedures are similar to Example 6. DNA  
bearing the forward zipper sequence (6-12 nt) on one end and the reverse zipper  
sequences (6-12 nt) on the other are digested briefly (.5-1.5 min) with exonuclease (T7  
gene 6 product is preferred). Vector-containing zipper sequences are digested in an  
10 identical manner. Ten-100 mg of vector and DNA insert so treated are annealed 20  
min-2 hr in 10 ml standard amplification buffer and used to transform competent *E. coli*  
by standard methods.

The cloning methods of Examples 6 and 7 have different advantages for  
15 particular experiments. In most cases, unmodified wild type vector is the preferred  
cloning vehicle when initiating a series of experiments and, as the addition of vector-  
containing ZAPs to insert DNA permits cloning into any unmodified vector to which  
sequence is known, the method of Example 6 is chosen. However, if an investigator  
wishes to repetitively use a particular vector in a long series of experiments, cloning  
20 into zipper-modified vector will be preferred (Example 7), as adapters need not be  
added to the PCR-generated DNA by either GLUEE or secondary PCR.

It should be noted that the GLUEE procedure can be used to modify vector  
DNA as well as PCR-generated DNA. This will provide another simple means to add  
25 zipper sequence to wild-type vector. It also demonstrates a means of adding special  
modified nucleotides and functional sequences to vectors without cloning

**Example 8.** *Release of zipper cDNA from vector and shuttle-cloning into other zipper-  
vector constructs*

DNA inserts in zipper-vectors have been constructed so that they can be released from plasmid with the zipper intact so that, after treatment with exonuclease, cohesive zipper sequence ends are regenerated and DNA can be joined to other ZAPs and cloned into other zipper vectors. A Pvu 1 site has been engineered at each junction of vector and zipper sequence (Fig. 11). This is not to be confused with the Apa 1 site at the zipper-insert junction. Thus, the order of sequences is vector, Pvu 1, zipper forward, Apa 1, insert, Apa 1, zipper reverse, Pvu 1, vector. Digestion of vector in which zipper inserts reside by Apa 1 leaves the zippers joined to vector (Example 7). Digestion with Pvu 1 releases insert with zippers intact (current example).

The purpose of regenerating zipper-cohesive DNA adds great flexibility to this system. Although DNA inserts could easily be reamplified for further recombinant experiments calling for modifications in sequence or chemical modifications of DNA ends, it is known that unwanted mutations can be introduced into DNA during amplification by Taq DNA polymerase. This may be especially relevant when isolating clones from cDNA libraries in zipper-modified lambda phage. In such cases, it will usually be necessary to subclone inserts into plasmids. Reamplification before subcloning may introduce errors that could distort DNA analysis. The ability to regenerate zipper-cohesive ends facilitates rapid shuttle subcloning without the risk of Taq-induced errors.

**Example 9. *Introduction of uridine-containing adapters into zipper-DNA for rapid cloning.***

One special application of procedures illustrated in Examples 3 and 4 is the use of GLUEE and secondary PCR to introduce ZAP-containing uridine. This procedure has been used to subclone portions of the human and baboon thrombin receptors, and G proteins G<sub>11</sub> and G<sub>14</sub> in this laboratory. The UDG cloning method exists in the art and has been shown to be highly effective. However, that method requires that many PCR primers be synthesized with uridine residues, which is inefficient and expensive. In the current invention, we have synthesized ZAP primers (U-FOR and U-REV) that bind to

cohesive zipper ends. After incorporation to DNA, the resulting construct can be digested with UNG according to the directions of the manufacturer (BRL). In one example, >700 colonies were obtained from 75 ng of 700 bp segment of human G<sub>11</sub> cDNA cloned by this method.

5

This example demonstrates that the ZAP methods can augment many protocols existing in the art. The key advantage of the invention is the reduction in number and types of primers that needs to be synthesized. Regardless of whether uridines, biotins, etc., need to be incorporated into a given DNA, such groups can be attached with a minimum of effort using the catalogued adapters.

10

**Example 10. *Introduction of zipper sequences into non-amplified DNA.***

Although many aspects of the current invention have been designed for practical use with PCR-generated DNA, it is evident that the invention will work well with any DNA to which the common-denominator 6-12 base zipper sequences have been added. Many methods exist in the art for such additions; genome DNA can be digested with restriction endonucleases to generate 2-6 base cohesive ends. Oligonucleotide adapters have been constructed so that the 6-12 base zipper sequences can be added to such DNA by standard ligation (Fig. 13) (SEQ ID NOS: 65-74). Blunt end adapters can also be used to adjoin sequences to blunt-end-digested genomic or cloned DNA.

15

20

Zipper sequences can also be added to cDNA during reverse transcription (RT) of mRNA using standard techniques, such that 6-12 bases can be added to the 5' of poly-dT oligonucleotide used to prime mRNA RT. Additional sequence can be added to the 5' of cDNA by tailing reactions or by ligation with RNA or DNA ligase.

25

Regardless, once defined zipper sequences have been added to nucleic acids, resulting molecules can be modified by ZAPs and cloned to zipper vectors. Guanine DNA cut by one restriction enzyme or produced by sonication will not be directionally manipulated by these procedures, as ligation will result in symmetrical placement of an

30

adapter at both ends. However, genomic DNA cut by two zippers, and cDNA reverse transcombined by zipper-oligo dT can be directionally cloned and manipulated. Again, it is evident that zipper methods detailed in this invention will serve as useful adjuncts to many protocols currently in the art. It is evident that cloning of both genomic DNA and cDNA will be greatly enhanced by extending cohesive ends for the currently  
5 available 2-4 bases generated by restriction enzymes to 6-12 bases by zipper methods.

**Example 11. *Enhanced long-distance sequencing of PCR products.***

10 Example 1 (Fig. 2) presents a simple method for selectively converting amplified DNA to single strands. It is evident that this invention can be used for extending and enhancing dideoxysequencing of PCR-amplified DNA. If at least 5 uridines are uncompartmented at the 5' of 1 primer but not the other in a pair of gene-specific primers used to amplify DNA, or joined to the DNA by GLUEE or secondary  
15 PCR, a highly stable 3' overhang can be created in the resulting PCR product by treatment with UNG enzyme (Fig. 14). This 3' overhang can protect that end of DNA from the 3' exonuclease activity of ExoIII enzyme. Addition of exonuclease III (ExoIII) to the PCR product selectively degrades the 3' of the other end. Fixed digestion of PCR product will result in progressive shortening of this end. After halting such reaction, the  
20 PCR strand can be reextended in the presence of fluorescent or radioactively-labeled nucleotides mixed with dideoxynucleotides, resulting in accumulation of DNA sequence information when said reactions are electrophoresed on acrylamide gel according to standard procedures.

25 Methods using the activity of ExoIII exist in the art for creating sets of nested clones that are useful in sequencing (as popularized by Strategene Corp., La Jolla, CA). In these procedures, protective 3' overhangs are created by restriction endonuclease digestion at one end of plasmid DNA, and ExoIII-susceptible 5' overhangs are created at the opposite DNA ends. These methods usually require that the progressively shortened  
30 DNA be subsequently "polished" with mung bean nuclease or similar endonuclease, circularized and cloned. Other investigators have directly sequenced ExoIII-treated

plasmids and cosmids after double restriction digestion and treatment with additional enzymes, including methalases. However, such procedures have not been easily applied to sequencing of PCR products, especially because 3' overhang-producing restriction enzymes do not easily cut the ends of PCR products. Thus, the merit of this example is that we have demonstrated that a highly reliable 3' overhang can be produced that selectively protects DNA ends against ExoIII digestion. After such a protective end is generated, we have found methods in the art highly useful for selectively sequencing long distances of amplified DNA.

**Example 12. Detection by the Point EXACCT Method**

**Cell Lines**

Purified DNA from 5 different cell lines with known point mutations was used for performing PCR. HL60 has the wild type DNA for codon 12 of the c-Ki-ras gene (GGT). A549 contains a homozygous serine mutation (AGT) and Calu-1 contains a cysteine mutation (TGT). Calu-1 is heterozygous at the c-Ki-ras locus, harbouring both a wild type and activated allele. SW480 contains a homozygous valine mutation (GTT) and SW1116 contains an alanine mutation (GCT) for base 2 of codon 12 of the c-Ki-ras gene.

**Reconstruction experiments**

Cell lines A549 and Calu-1 containing a mutation at base 1 of codon 12 of the c-Ki-ras gene were used for reconstruction experiments with the wild type cell line HL60. After culturing, cells were resuspended in 1 x phosphate buffered saline (PBS) and the concentration determined in a Burker count chamber. Subsequently, all cell lines were diluted, to a concentration of 2 million cells in 1 ml 1 x PBS. The Ki-ras codon 12 mutated cell lines were diluted, starting with 1 ml mutant cells in 1 x PBS (2 million cells) in 4 ml wild type cells in PBS. This 1:5 dilution was repeated several times resulting in a stepwise reduction of the number of cells mutated for base 1 of codon 12

of the c-Ki-ras gene used for reconstruction experiments with the wild type cell line HL60.

### DNA amplification

5

Cells of undiluted and diluted mixtures were isolated by the Qiaamp blood kit method and further used for Point-EXACCT. PCR was performed using 40 cycles (94°C, 30 Sec.; 55°C, 30 Sec.; 75°C, 40 Sec.) in an automated DNA Thermal Cycler (Perkin-Elmer Cetus). Purified genomic DNA (1 µg) was added to 80 µl of 10 mM Tris HCl (pH 8.3), 50 mM potassium chloride, 1.5 mM magnesium chloride, 250 pmole of each deoxyribonucleotide, 19.8 µmol/l of anti-sense primer, 28.1 µmol/l of sense primer, and 2.5 U of Taq DNA polymerase (Gibco BRL, Life Technologies, Inc.). A 204-bp region of the c-Ki-ras gene was amplified using oligonucleotides of the first exon of c-Ki-ras gene outside the codon 12 region: sense primer (5' TTT TTA TTA TAA GGC CTG CTG 3') (SEQ ID NO: 76) and anti-sense primer (5' TCA GAG AAA CCT TTA TCT GTA TCA AAG AAT GG 3') (SEQ ID NO: 77). The amplification products were subjected to electrophoresis in a 2% agarose gel and visualized by staining with ethidium bromide.

### 20 Point-EXACCT

After amplification with primers of the first exon of Ki-ras gene outside the codon 12 region, the products were supplemented with dithiotreitol (DTT) to a final concentration of 1 mmol/l and digested with 4U/µl T7 gene 6 exonuclease (United States Biochemical, Cleveland, Ohio) for 15 minutes at 37°C. The digested product was then heated to 75°C for 15 minutes to stop the reaction. Subsequently, 10 µl of digested amplification product was added to 38 µl hybridization buffer (4x SSC, 20 mM Hepes, 2 mM EDTA, 0.15% Tween 20), 1 µl (50 ng/µl) digoxigenin-labeled detection probe (5' pG TGG CGT AGG CAA GAG TGC CTT G-DIG-3') (SEQ ID NO:78) and 1 µl (50 ng/µl) of a biotinylated probe (5'-BIO-TA TAA ACT TGT GGT AGT TGG AGC TG - 3') (SEQ ID NO:79), containing the point mutation at the 3' side. This was

also done with probes directed against base 2 of codon 12 of the c-Ki-ras gene (see Table I). Separately, incubation was performed with four different biotinylated probes, containing the wild type (G) or point mutation (A,T,C) sequence for both bases 1 and 2 of codon 12 of the c-Ki-ras gene. The products were transferred to a 96- well flat-bottomed microtiter plate (Falcon) coated with 80  $\mu$ l biotinylated bovin serum albumin (BSA), (2  $\mu$ g/ml, Sigma Chemical Co.) in PBS for 1 hour at 37°C, washed 3 times with 100  $\mu$ l of PBS containing 0.05% Tween 20, and saturated with 50  $\mu$ l streptavidin (1 $\mu$ g-ml) with 0.5% gelatin/well for 1 hour. Amplification products were captured in the wells and after washing the plate 3 times with 100  $\mu$ l of PBS/Tween 0.05%, the products were ligated for 15 minutes with  $1.25 \times 10^{-3}$  Units T4 DNA ligase (Promega Corporation, Madison, USA) in 50  $\mu$ l 1 x ligase buffer (30 mM Tris-HCl, pH 7.8; 10 mM  $MgCl_2$ , 10 mM DTT, 0.5 mM ATP). A thermostable Taq DNA ligase (New England Biolabs) was also tested. The wells were then washed with 100  $\mu$ l PBS/0.5% Tween 20 and denatured with 50  $\mu$ l 0.07 M NaOH for 2 minutes, washed twice with 100  $\mu$ l 0.01 M NaOH/0.05% Tween 20 and once with PBS/Tween 0.05%. The products were incubated with 8  $\mu$ l/well of a 1:2000 dilution of horse-radish peroxidase (HRP)-conjugated anti-digoxigenin (DIG) antibody Fab fragments (Boehringer Mannheim Biochemicals, Indianapolis, USA) in 100 mM HCl, pH 7.5/150 mM NaCl supplemented with 1% blocking reagent (Boehringer Mannheim Biochemicals, Indianapolis, USA) for 1 hour and washed 3 times with 100  $\mu$ l PBS/Tween 0.05%. After washing, 100  $\mu$ l/well of a 3,3',5,5'-tetramethyl-benzidine (TMB) chromogen solution (10 mg/ml) was added and the color development stopped with 50  $\mu$ l/well 2 M phosphoric acid. Previous steps in microtiter plate were performed at room temperature with shaking. The plates were read at 450 nm in a BIO-RAD Novapath Microplate Reader.

The method of Point-EXACCT is demonstrated in Figure 16. For example the method was used with cell line Calu-1, containing both wild type and activated allele for base 1 of codon 12 of the c-Ki-ras gene. PCR for Ki-ras was performed for amplification of a 204-bp product of exon 1 of the c-Ki-ras gene using primers outside the codon 12 region (Table I). These products were incubated with the T7 gene 7



exonuclease. This non-processive 5' exonuclease digests double-stranded DNA to single strands; digestion of each strand is expected to proceed at equivalent rates and, as the enzyme is inactive against single-stranded DNA, the reaction terminates at the approximate midpoint of the amplification product, leaving two single-strand DNAs of approximately half original product length. As an alternative, exonuclease digestion can be blocked on one strand by enzyme phosphorothioate linkages into one amplification primer. Previously, 3'-5' phosphorothiate linkages have been shown sufficient to inhibit exonuclease digestion. The digested products are separately hybridized with one common digoxigenin labeled probe and four different biotinylated capture probes containing the wild type of point mutation sequence for base 1 of codon 12 of the c-Ki-ras gene (Table I). After hybridization, the product-probe hybrids are captured in microtiter wells coated with biotinylated BSA and saturated with streptavidin. After capturing and washing, the products were incubated with T4 DNA ligase. T4 DNA ligase catalyzes the formation of a phosphodiester linkage between the two adjacent oligonucleotides, the 5' biotinylated probe containing the point mutation at the 3' side and the 3' labeled detection probe. If there is perfect complementarity, the enzyme T4 DNA ligase covalently joins the 5' biotinylated probe and the 3' labeled detection probe. If the probes and target are mismatched at their junction, a covalent bond is not formed. After ligation, denaturation and subsequent washing, amplification products are detected by horse radish peroxidase-conjugated anti-digoxigenin-antibody.

### *Coating*

Several coating substrates were used for capturing the amplification products in a 96-well plate: direct coating with avidin or streptavidin or coating with biotinylated BSA saturated with streptavidin. The best signal/noise ratio was measured for these substrates. Coating with avidin gave a very strong signal, but high cross-reactivity. Direct coating with streptavidin was less efficient compared to coating with biotinylated BSA, saturated with streptavidin (data not shown). Capturing the amplification products on a 96-well plate coated with biotinylated BSA and saturated with streptavidin gave the best results (data not shown).

*T7 gene 6 exonuclease*

A serial dilution of the enzyme T7 gene 6 exonuclease was used for efficiency of exonuclease conversion to single-stranded DNA. Small amounts of the enzyme (0.4 U/ $\mu$ l) effectively converted the amplification products into single strands, but incubation with a higher amount of the enzyme (4 U/ $\mu$ l) gave completely digested amplification products. These products digested with a high exonuclease concentration gave the best results at the capturing procedure. The hybridization efficiency after heat-denaturation alone, thus without exonuclease digestion was smaller compared to exonuclease-treated products, i.e. 60-70% of relative optical density obtained after exonuclease-treated products for undiluted products. This is reasonably convincing that exonuclease pretreatment is especially useful for detection of low frequency signals.

*T4 DNA ligase activity*

The most critical parameter of the procedure appeared to be the ratio DNA/ligase concentration. Therefore, products were incubated with a serial dilution of T4 DNA ligase varying from  $1.25 \times 10^{-1}$  U to  $1.25 \times 10^{-4}$  U for each digested product. The use of a high ligase concentration resulted in non-specific reactivity because a higher signal with control probes was found. First, maximal extinction was obtained with incubation of  $1.66 \times 10^{-2}$ , but after optimization of other parameters, maximal efficiency was obtained with incubation of  $1.25 \times 10^{-3}$  U T4 DNA ligase. Detection of the amplification products ligated with  $4 \times 10^{-3}$  U T4 DNA ligase. Detection of the amplification products ligated with  $4 \times 10^{-3}$  U to  $8 \times 10^{-3}$  U resulted in low sensitivity.

For optimizing the enzyme temperature and incubation time, the following conditions were tested for maximal sensitivity and specificity: room temperature for 15 minutes, 37°C for 15 minutes, 37°C for 15 minutes with addition of 0.5 mM ATP, 16°C overnight and 4°C overnight. Incubation of T4 DNA ligase at 37°C for 15 minutes with addition of 0.5 mM ATP gave non-specific ligation. Maximal efficiency was obtained with incubation of  $1.25 \times 10^{-3}$  U T4 DNA ligase for 15 minutes at room temperature.

Furthermore, inclusion of formamine can increase specificity of T4 DNA ligase. A higher yield was not found after incubation of T4 DNA ligase overnight at 4°C or overnight at 16°C. Incubation of a thermostable Taq DNA ligase did not improve sensitivity or specificity.

5

#### *Anti-digoxigenin antibody dilution*

To test the optimal staining procedure, several dilutions of one PCR reaction were tested. Initially, a 1:1000 dilution of alkaline phosphatase-conjugated anti-digoxigenin Fab fragments in PBS/0.5% Tween/0.5% BSA was used for detection with o-phenylenediamine dihydrochloride. With this dilution of anti-digoxigenin, a 1:100 dilution of the amplification products could specifically be detected. The sensitivity was improved to 1:1000 diluted amplification product by using a 1:1000 dilution of horse-radish peroxidase-conjugated anti-digoxigenin-Fab fragments in PBS/0.5% Tween/0.5% BSA or in hybridization mix, and detection with TMB. However, the best sensitivity, more than 1:10,000 diluted amplification product was obtained when using a 1:2000 dilution of horse-radish peroxidase-conjugated anti-digoxigenin-Fab fragments in 100 mM TrisHCl, pH 7.54/150 mM NaCl supplemented with 1% blocking reagent of 5% heat inactivated fetal calf serum.

20

#### *Washing*

Washing procedures must be carefully carried out, reducing the risk of carry over. After eliminating the washing substrate from one well, the pipette is washed with another volume PBS/0.05% Tween 20, reducing the risk of carry over.

25

#### **Analysis of the cell lines for Point-EXACCT**

Five different cell lines with known point mutations at base 1 and 2 of codon 12 of the c-Ki-ras gene were used for detection of Ki-ras point mutations with Point-EXACCT (Table II). The results of Point-EXACCT in the Ki-ras model system are

30

shown in **Figure 17**. Maximal extinctions were obtained from hybridized samples without ligation. The whole experiment, including DNA extraction and Polymerase Chain Reactions, takes about 10 hours.

## 5 **Reconstruction experiments**

The sensitivity of the Point-EXACCT procedure was tested by reconstruction experiments. Therefore, at least two methods are available. First, extracted DNA of the mutant cell line can be mixed with extracted DNA of the wild type cell line. However,  
10 this method does not give a good quantitative measurement of the cells because of mixing DNA. Therefore, a reconstruction method was used by mixing vital mutant cells with vital wild type cells. For this experiment a serial dilution (1:5) was made for both mutant cell lines A549 and Calu-1, and HL60, leading to a reduction in the ratio mutant cells to wild type cells. Next, DNA of mixed cell suspensions was extracted.  
15 This method gives a good quantitative measurement since the number of cells is known.

Each cell mixture of the different mutant cell lines Calu-1 and A549 and wild type cell line HL60 was amplified and further tested with Point-EXACCT. The results of the reconstruction experiments of mutant cell lines Calu-1 and A549 with the wild  
20 type cell line HL60 are shown in **Figures 18A and B**. Note that 3 x standard deviation plus the mean (represented by the broken line) equals a 10.0 and 8.1 relative optical density unites for mutant cell line A549 and mutant cell line Calu-1, respectively. For these values, a chance of 0.5% is present this value belongs to a normal population. For the 1:78,125 dilutions the relative optical densities is along 5% (mean + standard  
25 deviation). Therefore, these values can safely be regarded as detectable and significantly different from normal.

The present study describes the vulnerable steps and the sensitivity of the Point-EXACCT procedures. The method of Point-EXACCT is capable of selective detection  
30 of point mutations in base 1 and base 2 of codon 12 of the c-Ki-ras gene. **Figure 17** shows the relative optical densities of the different cell lines after incubation with 4

biotinylated probes, containing the wild type (G) and point mutation (A, T, C) sequence for base 1 and base 2 of codon 12 of the c-Ki-ras gene. The probe detecting the wild type for the first base of Ki-ras codon 12 showed only signals comparable to backgrounds in the cell lines with point mutations. Cell line HL60 showed only a  
5 signal after incubation with the wild type probe. The GGT to AGT mutation (glycineserine) was only detected in amplified DNA of cell line A549. In Calu-1, both wild type and the mutant for cysteine coding (GGT to TGT) were detected. No cell line containing the GGT to CGT mutation was tested with the mutation specific probes. However, incubation of the mutation specific probe CGT gave no signal above the  
10 background values using DNA of either cell line HL60, A549 or Calu-1 (Figure 17). Two cell lines SW480 and SW1116 were tested for the detection of point mutations at base 2 of codon 12 of the c-Ki-ras gene. The GGT to GTT mutation (glycine-valine) was detected in cell line SW480 and both wild type and activated allele (GGT and GCT) were detected in cell line SW1116. Incubation with control probes gave no  
15 signal above background values (Figure 17).

Several parameters for the Point-EXACCT method were optimized to obtain a detailed understanding of the various parameters influencing the Point-EXACCT method. For additional detail on these techniques, see references 43-47. The most  
20 important parameters for high sensitivity and absence of unspecific background of the Point-EXACCT method were: the optimal coating substrate, the concentration enzyme T7 gene 6 exonuclease, the concentration T4 DNA ligase, the use of a blocking reagent during immunochemical detection, the concentration of anti-digoxigenin-antibody Fab fragments, the proper selection of the marker enzyme catalyzing a highly sensitive color  
25 reaction, and the washing procedures.

The optimal coating was the most important parameter for capturing the amplification products in a 96-well plate. Coating with biotinylated bovin serum albumin and saturation with streptavidin gave the best results. Other coating substrates  
30 resulted in non-specific reactivity.

The sensitivity of the detection of the amplified fragments was influenced by using the enzyme T7 gene 6 exonuclease. Digestion of a double-stranded amplification product to single strands increased hybridization efficiency and conferred increased sensitivity and specificity of detection by eliminating competition of complementary DNA with the oligonucleotide probes. T7 gene 6 exonuclease is active in the amplification reaction buffer as recommended by the manufacturer when supplemented with dithiotreitol, so that amplification, exonuclease digestion, and hybridization can quickly and sequentially be done in the same reaction tube prior to transfer to the microtiter plate. The inclusion of dithiotreitol improved the efficiency of T7 gene 6 exonuclease. Digestion of the amplified products with T7 gene 6 exonuclease gave a greater efficiency compared to a heat-denatured amplification product. This is explained by the possibility of reannealing of undigested denatured amplification products and, therefore, reduces the efficiency of hybridization with biotin and digoxigenin labeled probes.

Furthermore, the concentration of T4 DNA ligase appeared to be the most critical parameter of the Point-EXACCT method. The T4 DNA ligase concentration, rather than the incubation temperature or time, yielded the best results. Optimal ligase efficiency was observed with incubation of  $1.25 \times 10^{-3}$  U T4 DNA ligase. This concentration gave the best signal/noise ratio. Reduction of ligase concentration from  $10^{-3}$  U to  $8 \times 10^{-3}$  U resulted in reduced sensitivity. Using a high T4 DNA ligase concentration resulted in non-specific ligation. An important component of the anti-digoxigenin-antibody Fab fragment solution is the reagent used for blocking of the 96-well plate against unspecific binding of anti-digoxigenin-antibody Fab fragments conjugated with horse-radish peroxidase during the detection reaction. This blocking reagent contains fractionated milk proteins with small quantities of carbohydrates and fructosamines, having a reduction potential which may explain a lower background.

Concerning the selection of the marker enzyme and staining for the final detection step, horse-radish peroxidase with 3,3',5,5'-tetramethyl-benzidine was favoured over alkaline phosphatase with o-phenylenediamine dihydrochloride.

The final parameter influencing the Point-EXACCT procedure are the washing procedures. Because the Point-EXACCT method is highly sensitive, washing procedures should be carefully carried out to avoid carryover. The invention contemplates that removing uncaptured labeled probe may be performed with a variety of detergent wash compositions, including for example TWEEN solution.

The present concentrations of various parameters, especially the use of exonuclease and the high concentration of biotin and digoxigenin probes allow a sensitive detection. In case a less sensitive detection is desired, the concentrations of the above-mentioned parameters may be lowered. The invention contemplates that the capture to solid support may be performed either before, or after, the denaturing of the nucleic acid strands, or before, or after, ligating the probes. Additionally, the invention contemplates that both the first and second probes may be adapted with detectable moieties, rather than a capture moiety, such that ligation can be detected by a differential energy transfer detection assay.

After optimizing the different parameters influencing the Point-EXACCT method, the surprising sensitivity was shown by reconstruction experiments (**Figure 20A and 20B**). These experiments demonstrate the detection of 1 mutant cell in more than 75,000 wild type cells. Further dilution (1:380,000) gave no reliable signal above the background values. Point-EXACCT here used a 8.5 microliter amplification product after a PCR reaction containing 4 micrograms DNA in 100 microliters amplification reaction buffer. Therefore, the number of 78,125 cells may not be reached theoretically since 1 cell contains 5-7 pg DNA. However, there is a chance that one mutant cell is present in the amplification tube by which a signal could be detected with Point-EXACCT, represented in **Figure 18A**. This was not the case with cell line Calu-1, where only 1 mutant cell in 15,000 wild type cells could be detected (**Figure 18B**). There was a decrease in the relative optical density obtained after hybridization with the probe directed against the specific mutation with increasing dilutions of the mutant cells to wild type cells. The reconstruction experiments were repeated using new cell cultures, counting, mixing and amplification of the cell mixtures. Those

experiments showed again the detection of 1 mutant cell in an excess of 78,125 and 15,000 wild type cells in cell line A549 and Calu-1, respectively, using 4 micrograms DNA in 100 microliters amplification reaction buffer.

5           The Point-EXACCT method is also surprisingly rapid. Starting from genomic DNA, or alternatively from paraffin or frozen sections, blood cells, bronchial epithelial cells obtained during bronchoscopy or sputum, as a source of template DNA, the detection of Ki-ras point mutations can be accomplished within 10 hours. Analysis of point mutation detection by this method requires considerably less time or effort than  
10 other techniques used for the detection of point mutations since this procedure can be performed in one day. For performing the Point-EXACCT method, only small numbers of cells or DNA samples are sufficient for analysis. Therefore, this method is useful for microdissection studies. Furthermore, the method is not harmful since it does not require the use of radioactive materials. Essential in this method is that only DNA  
15 molecules which are hybridized with two adjacent probes, which are subsequently ligated, will be detected by this format, giving a very high degree of specificity. T4 DNA ligase catalyzes the formation of a phosphodiester linkage between two adjacent oligonucleotides, the 5' biotinylated probe containing the point mutation at the 3' side and the 3' labeled detection probe: if there is perfect complementarity, the enzyme T4  
20 DNA ligase covalently joins the 5' biotinylated probe and the 3' labeled detection probe. If the probes and target are mismatched at their junction, a covalent bond is not formed.

Furthermore, the invention contemplates a variety of useful combinations of adjacent probes. For example, a first probe may be adapted with a capturable moiety on  
25 the 3' end and a second probe may be labeled with a detectable moiety on the 5' end, such that ligation is performed between the 5' end of the first probe and the 3' end of the second probe. Alternate configurations are also contemplated by the present invention, such that the mutation can be on the 5' side of a 3' labeled probe, for example.

30           In addition to its simplicity, speed, specificity and sensitivity, the Point-EXACCT method generates quantitative output (Figure 18A and 18B), providing and



opportunity to measure the actual fraction of the sample DNA carrying a point mutation. The Point-EXACCT method can be used for the detection of early occurring point mutations, such as ras and other oncogenic mutations, in a heterogeneous cell population, where only a small fraction of tumor cells containing a mutation are present

5 among much larger numbers of genetically normal cells. See, Oncogenes, K.B. Burck, et al., Springer Varley & Co, N.Y. (1988); or The Oncogene Handbook, T. Curran et al., Elsevier Science Publ., Netherlands (1988). Other genetic disorders caused by point mutations can be diagnosed by the present invention, such as cystic fibrosis, for example. The simplicity, accuracy, and reproducibility of Point-EXACCT makes it

10 suitable for screening large populations for, presence of known point mutations, deletions, insertions or polymorphisms.

**TABLE I.** Ki-ras codon 12 5' mutation specific biotinylated oligonucleotide probes and 3' digoxigenin labeled probes used for detection of point mutations in K-ras gene

(a) Ki-ras codon 12 biotin-labeled mutation specific capture probes (the specific mutation is underlined)

5

<u>Probe</u>	<u>Wild type/Mutation</u>	<u>Sequence</u>
<i>base 1</i>		
10 K12G	GGT	5'BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TG</u> - 3' (SEQ ID NO:79)
K12A	GGT to AGT	5'BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TA</u> - 3' (SEQ ID NO:80)
K12T	GGT to TGT	5'BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TT</u> - 3' (SEQ ID NO:81)
15 K12C	GGT to CGT	5' BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TC</u> - 3' (SEQ ID NO:82)
<i>base 2</i>		
20 K12G	GGT	5' BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TGG</u> - 3' (SEQ ID NO:83)
K12A	GGT to GAT	5' BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TGA</u> - 3' (SEQ ID NO:84)
K12T	GGT to GTT	5' BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TCC</u> - 3' (SEQ ID NO:85)
25 K12C	GGT to GCT	5' BIO - TA TAA ACT TGT GGT AGT TGG AGC <u>TGT</u> - 3' (SEQ ID NO:86)

(b) K-ras 3' digoxigenin labeled detection probes

30

<u>Probe</u>	<u>Sequence</u>
<i>base 1</i>	
K-ras DIG	5' - pG TGG CGT AGG CAA GAG TGC CTT G - DIG 3' (SEQ ID NO:78)
<i>base 2</i>	
K-ras DIG	5' - pTGG CGT AGG CAA GAG TGC CTT G - DIG 3' (SEQ ID NO:87)

35

40

**TABLE II.** Different cell lines used for detection of point mutations in base 1 and 2 of codon 12 of the c-Ki-ras gene

	<u>Cell line</u>	<u>Sequence</u>	<u>Amino Acid</u>
5	HL60	GGT	glycine
	A549	AGT	serine
	Calu-1	TGT	cysteine
	SW1116	GCT	alanine
10	SW480	GTT	valine

\* \* \* \* \*

From the foregoing, it can be seen that, although specific embodiments of the invention have been described as a means of illustration, many different variations can be made and proposed without deviating from the spirit and scope of the invention. The content of the references cited herein are hereby incorporated by reference in their entireties.

## REFERENCES

1. Saiki, RK, Walsh DS, Erlich, HA. *Proc. Natl. Acad. Sci. USA* 1989;86:6230-4
2. White TJ, Madej, R, Persing DH. *Adv. Clin. Chem.* 1992;29:161-96
3. Persing DH. *J. Clin. Microbiol.* 1991;29:1281-5
4. Maniatis et al. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1983
5. Jung V, Pestka SB, Pestka S. *Nucleic Acid Res* 18:6156
6. Saiki RK, Walsh DS, Erlich HA. *Proc Natl Acad Sci USA* 1989;86::6230-4
7. Zhang Y, Coyne MY, Will SG, Levenson CH, Kawasaki ES. *Nucleic Acids Res.* 1991;19:3929-33
8. Helmuth R. Nonisotopic detection of PCR products. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds.), *PCR Protocols: A Guide to Methods and Applications*. San Diego, CA: Academic Press, pp. 119-28, 1990.
9. Gyllensten UB, Erlich HA. *Proc Natl Acad Sci U S A* 1988;85:7652-56
10. Stoker AW. *Nucleic Acids Res* 1990;13:4290
11. Aslanidis C, de Jong P. *Nucleic Acids Res* 1990;18:6069-74
12. Kuijper *Gene* 1992;112:147-55
13. Haun RS et al. *BioTechniques* 1992;13:515-8
14. Haun R, Moss J. *Gene* 1992;112:37-43
15. Pham T, Cheng QL. *CLONTECHques* 1993;8:5-9
16. Kaluz S, Kolble K, Reid KB. *Nucleic Acids Res* 1992;20:4369-70
17. Lindahl T, Ljungquist S, Siebert W, Nyberg B, Sperens B. *J Biol Chem* 1977;252:3286-94

18. Rashtchian A, Buchman GW, Schuster DM, Berninger MS. *Anal Biochem* 1992;206:91-7
19. Rashtchian A, Thornton CG, Heidecker G. *PCR Methods Appl* 1992;2:124-30
20. Nisson PE, Rashtchian A, Watkins PC. *PCR Methods Appl* 1991;1:120-3
21. Mitchell and Merrill *Anal Biochem* 1989;178:239-42
22. Uhlen *Proc Natl Acad Sci USA* 1990;87:6569-73
23. Kerr, C, Sadowski PD. *J Biol Chem* 1972;47:311-8
24. Straus, N.A., & Zagursky, R.J. *Biotechniques* 1991;10:376-84
25. Anderson LJ, Gary WG, Young N. Human parvovirus B19. In Lennette ED (ed.) *Laboratory Diagnosis of Viral Infections*. Marcel Dekker, New York, pp. 627-42, 1992
26. Török TJ. Parvovirus B19 and human disease. *Adv Intern Med* 1992;37:431-455
27. Durigon EL, Erdman DD, Gary GW, Pallansch MP, Torok TJ, Anderson JJ. *J. Virol. Methods* (in press).
28. Török TJ, Wang QY, Gary GW, Yang CF, Finch TM, Anderson LJ. *Clin Infect Dis* 1992;14:49-53
29. de Noronha CM, Mullins JJ. *PCR Methods Appl* 1992;2:131-6
30. Eisenach KD, Sifford MD, Cave MD, Bates JH, Crawford JT. *Ann Rev Respir Dis* 1991; 144:1160-3
31. Böddinghaus B, Rogall T, Flohr T, Blöcker H, Böttger EC. *J Clin Microbiol* 1990;28: 1942-6
32. Woese CR. *Microbiol Rev* 1987;51:221-71
33. Wilson KH, Blitchington RB, Greene RC. *J Clin Microbiol* 1990;28:1942-6
34. Roberts GD, Goodman NL, Heifets L, Larsh HW, Lindner TH, McClatchy JK, McGinnis MR, Siddiqi SH, Wright P. *J Clin Microbiol* 1983;18:689-96

35. Drake T, Hindler JA, Barton GW, Bruckner DA. *J Clin Microbiol* 1987;25:1442-5
36. Evans KD, Nakasome AS, Sutherland PA, DeLaMaza LM, Peterson EM. *J Clin Microbiol* 1992;30:2472-31
37. Jarolim P et al. *Proc Natl Acad Sci USA* 1991;88:11022-6
38. Guo LH, Wu R. *Nucleic Acids Res* 1982;10:2065-84
39. Guo LH, Wu R. *Methods Enzymol* 1983;100:60-9
40. Henikoff S. *Gene* 1984;28:351-9
41. Sorge JA, Blinderman LA. *Proc Natl Acad Sci USA* 1989;86:9208-12
42. Li C, Tucker PW. *Nucleic Acids Res* 1993;21:1239-44
43. Saiki, K. et al. *Science* 1988;239:487-490
44. Holloway, B. et al. *Nucleic Acids Res.* 1993;21:3905-3906
45. Bos, J.L. *Mut. Res.* 1988;195:255-271
46. Valenzuela, D. and Groffen, J. *Nucleic Acids Res.* 1985;14:843-851
47. Capon, D.J. et al. *Nature* 1983;304:507-513

65

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Murtagh, James J.
- (ii) TITLE OF INVENTION: METHODS FOR NUCLEIC ACID DETECTION,  
SEQUENCING AND CLONING USING EXONUCLEASE
- (iii) NUMBER OF SEQUENCES: 87
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: NEEDLE & ROSENBERG, P.C.
  - (B) STREET: Suite 1200, The Candler Bldg., 127  
Peachtree Street N.E.
  - (C) CITY: Atlanta
  - (D) STATE: Georgia
  - (E) COUNTRY: USA
  - (F) ZIP: 30303
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: U.S. Serial No. 08/479,723
  - (B) FILING DATE: U.S. Filing Date: 07-JUN-1995
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Perryman, David G.
  - (B) REGISTRATION NUMBER: 33,438
  - (C) REFERENCE/DOCKET NUMBER: 13013.0002/P
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (404) 688-0770
  - (B) TELEFAX: (404) 688-9880

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 12 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGAGGGAAGA GG

12

66

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGCACGCGGG AG

12

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGTAATTCCG GACAACGCTC

20

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTCACGAAC AACGCGACAA

20

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid



67

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACGGTGCCCG CAAAGTGTGG

20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ACCGTGAGGG CATCGAGGTG

20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCCCGCCGC GCCCGCGCC CGGCCCGCCG CCCCCGCCGG GCCCGAGGGA AGAGG

55

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGGAGACCG GAATTCGGAT CCCATATGGC GGCCGCGATC GAGGGAAGAG G

51

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GACAGAGCAC AGAATTCGGA TCCAAGCTTC GATCGCACGC GGGAG

45

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CAUCAUAC AUCAUCAUGC GATCGAGGGA AGAGG

35

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CUACUACUAC UACGATCGCA CGCGGGAG

28

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 53 base pairs

69

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CATCATCACA GCAGCGGCTT GGTGCCGCGC GGCGGCGCGA TCGAGGGAAG AGG

53

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCAGCTTCCT TTCGGGCTTT GTTAGCAGCC GGATCCGATC GCACGCGGGA G

51

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AACATGAGA GGATCGCATC ACCATCACCA TCACGGATCC TCGATCGAGG GAAGAGG

57

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGATCTATCA ACAGGAGTCC AAGCTCAGCT AATTAAGCTT CGATCGCACG CGGGAG 56

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGTAAAACGA CGGCCAGTGA ATTGTAATAC GACTCACTAT AGGGCGAATT CGATCGAGGG 60  
AAGAGG 66

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AACAGCTATG ACCATGATTA CGCCAAGCTA TTTAGGTGAC ACTATAGAAT ACTCAAGCTT 60  
GGATCCGATC GCACGCGGGA G 81

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGAGGGAAGA GGTACAGCAT CATCTTTGTG GTGGGGA 37

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGCACGCGGG AGTCGAATTC CGAGACTCAT AATGA

35

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGAGGGAAGA GGTAGGTCAG CAACATGGAC TATGT

35

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCACGCGGG AGTAGCGGGC CAAGGCGAAC C

31

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

72

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGAGGGAAGA GGTGCCTCCC CAACAAGACT GCCA

34

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGCACGCGGG AGTCCACATG TCTCCCAGCA GATG

34

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGAGGGAAGA GGTCCATGGA GAAGGCTGGG G

31

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCACGCGGG AGTCAAAGTT GTCATGGATG ACC

33

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CGAGGGAAGA GGTCCNYTNY TNYTNGGNGC NGCNATGGC

39

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CGCACGCGGG AGTANACCCA NGGRTCWARD ATYTGRITCC A

41

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGCCCGCCGC GCCCGCGCC C

21

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

74

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGCACGCGGG AGTANARNGC RAANGGRTTN ACRCA

35

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGAGGGAAGA GGTAYGGNG ARATHGGNAT HGGNACNCC

39

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGCACGCGGG AGTYTGNGTN ACNGTDATNC CNCCNACNGT

40

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CGAGGGAAGA GGTGCNGGNG ARWSNGGNAA RWSNAC

36



## (2) INFORMATION FOR SEQ ID NO:33:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CGCACGCGGG AGTCNSWNCK YTGNCNACN ACRTC

35

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CGAGGGAAGA GGATGGGATG TACTCTGAGC GCAGAGG

37

## (2) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CGCACGCGGG AGCGCACGCG GGAGTCAAAG TTGTCATGGA TGACC

45

## (2) INFORMATION FOR SEQ ID NO:36:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

76

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CGAGGGAAGA GGATGACTCT GGAGTCCATC ATGGCGT

37

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CGCACGCGGG AGTCCATGGA GAAGGCTGGG GCC

33

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CGAGGGAAGA GGATGGCCGG CTGCTGCTGT TTGTCTG

37

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

77

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CGAGGGAAGA GGATGACTCT GGAGTCCATG ATGGCGT

37

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CTCCCGCGTG CGATCGGTCA TAGCTGTTTC CTGTG

35

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CACAGGAAAC AGCTATGACC GATCGCACGC GGGAG

35

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTCCCGCGTG CGATCGGGCT GCTGCCACCG CTGAGC

36

(2) INFORMATION FOR SEQ ID NO:43:

78

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GCTCAGCGGT GGCAGCAGCC CGATCGCACG CGGGAG

36

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CGGCACCGTC ACCCTGGATG CTGTAGG

27

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CGAGGGAAGA GGTGGCGATT AAGTTGGGTA ACGCC

35

## (2) INFORMATION FOR SEQ ID NO:46:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGCACGCGGG AGTCACACAG GAAACAGCTA TGAC

34

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGAGGGAAGA GGTGTGTGGA ATTGTGAGCG G

31

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GTAAAACGAC GGCCAGT

17

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CGCCAGGGTT TTCCAGTCA CGAC

24

(2) INFORMATION FOR SEQ ID NO:50:

80

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GAGCGGATAA CAATTCACA CAGG

24

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CGAGGGAAGA GGTGGCGATT AAGTTGGGTA ACGCC

35

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GGCAGCGGG AGTCACACAG GAAACAGCTA TGAC

34

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid

81

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CGAGGGAAGA GGTGTGTGGA ATTGTGAGCG G

31

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CGAGGGAAGA GGGGCCCCC CCCCCGGGT TTAAAAAAA AAAA

44

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CGACGTTTTT TTTTTTAA CCGGGGGGG GGGGGCC

37

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGAGGGAAGA GGTGGCACGC GGGAGT

26

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TGTAAAACGA CGGCCAGTAG CAAGTTCAGC CTGGTTAA

38

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCCCAGTCAC GACGTTGTGG GGTAAATAAC AGAGGTGGC

39

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GGTGGCGACG ACTCCTGGAG CCCG

24

(2) INFORMATION FOR SEQ ID NO:60:



83

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TTGACACCAG ACCAACTGGT AATG

24

## (2) INFORMATION FOR SEQ ID NO:61:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CGACGGCCAG TCGACTCTAG TTTTTTTTTT TT

32

## (2) INFORMATION FOR SEQ ID NO:62:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CGACGGCCAG TCGACTC

17

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GATCCATCGA GGGAAGAGGG GCCCTCTCCC GCGTGCCA

38

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GTAGCTCCCT TCTCCCCGGG AGAGGGCGCA CGGTTCCA

38

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CGAGGGAAGA GGTGCGGCCG CTT

23

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CGCCGGCGGAA

10

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CGAGGGAAGA GGTGCGGCCG C

21

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CGCCGGCGGTT AA

12

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CGCCGGCGGCT AG

12

86

## (2) INFORMATION FOR SEQ ID NO:70:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CGCCGGCGTC GA

12

## (2) INFORMATION FOR SEQ ID NO:71:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GCGGCCGC

8

## (2) INFORMATION FOR SEQ ID NO:72:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

AATTGCGGCC GC

12

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid

87

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GATCGCGGCC GC

12

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

AGCTGCGGCC GC

12

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

AAGCGGCCGC

10

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

TTTTTATTAT AAGGCCTGCT G

21

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TCAGAGAAAC CTTTATCTGT ATCAAAGAAT GG

32

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

GTGGCGTAGG CAAGAGTGCC TTG

23

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

TATAAACTTG TGGTAGTTGG AGCTG

25

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

89

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TATAAACTTG TGGTAGTTGG AGCTA

25

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

TATAAACTTG TGGTAGTTGG AGCTT

25

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

TATAAACTTG TGGTAGTTGG AGCTC

25

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

TATAAACTTG TGGTAGTTGG AGCTGG

26

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TATAAACTTG TGGTAGTTGG AGCTGA

26

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TATAAACTTG TGGTAGTTGG AGCTCC

26

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

TATAAACTTG TGGTAGTTGG AGCTGT

26

(2) INFORMATION FOR SEQ ID NO:87:



(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

TGGCGTAGGC AAGAGTGCCT TG

23

## WHAT IS CLAIMED IS:

1. A method of detecting the presence of a nucleotide sequence within a double-stranded DNA in a sample comprising:
  - a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA;
  - b. hybridizing the single-stranded DNA with i) a first nucleic acid probe adapted with a moiety which can be captured to a solid support, and ii) a second nucleic acid probe labeled with a detectable moiety which can hybridize with the single-stranded DNA adjacent the hybridized first nucleic acid probe;
  - c. ligating the hybridized first and second nucleic acid probes;
  - d. capturing the moiety on the first nucleic acid probe hybridized to the DNA to the solid support;
  - e. denaturing the ligated first and second nucleic acid probes from the hybridized single-stranded DNA;
  - f. removing uncaptured labeled probe; and,
  - g. detecting the presence of captured detectable moiety, the presence of captured detectable moiety indicating the presence of the nucleotide sequence within the double-stranded DNA in the sample.
2. The method of claim 1, wherein the exonuclease is selected from the group consisting of T7 gene 6, exonuclease III, T4 DNA polymerase, non-recombinant T7 DNA polymerase, and lambda exonuclease.
3. The method of claim 1, wherein the double-stranded DNA is contacted with an exonuclease for a time sufficient to convert the entire double-stranded DNA to single-stranded DNA.
4. The method of claim 1, wherein the moiety on the first nucleic acid probe is biotin, and the capturing step is performed by capturing the biotin on a streptavidin-coated solid support.

5. The method of claim 1, wherein the detectable moiety on the second nucleic acid probe is digoxigenin, and the detecting step is performed by binding the digoxigenin with anti-digoxigenin Fab fragments.
6. The method of claim 5, wherein the detecting step is performed using about a 1:2000 dilution of horse-radish peroxidase-conjugated anti-digoxigenin Fab fragments in about 100mM Tris HCl, pH 7.5 in 150 mM NaCl supplemented with 1% blocking reagent.
7. The method of claim 1, wherein the removing step is performed by washing the sample with a detergent solution.
8. The method of claim 1, wherein the ligating step is performed with  $1.25 \times 10^{-3}$  U T4 DNA ligase for fifteen minutes at room temperature.
9. The method of claim 8, wherein the ligating step is further performed with formamide.
10. A method of detecting a point mutation within a double-stranded DNA in a sample comprising:
  - a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA;
  - b. hybridizing the single-stranded DNA with i) a first nucleic acid probe containing the point mutation at the 3' end and adapted with a moiety at the 5' end which can be captured to a solid support, and ii) a second nucleic acid probe labeled with a detectable moiety at the 3' end and which can hybridize at the 5' end with the single-stranded DNA adjacent the hybridized 3' end of the first nucleic acid probe;
  - c. ligating the hybridized first and second nucleic acid probes;
  - d. capturing the moiety on the first nucleic acid probe hybridized to the DNA to the solid support;

- e. denaturing the ligated first and second nucleic acid probes from the hybridized single-stranded DNA;
- f. removing uncaptured labeled probe; and,
- g. detecting the presence of captured detectable moiety, the presence of captured detectable moiety indicating the presence of the point mutation within the double-stranded DNA in the sample.

11. A method of detecting a point mutation within a double-stranded DNA at a level of one copy of the point mutation per at least about 75,000 normal copies in the sample, comprising:

- a. digesting the double-stranded DNA with an exonuclease which converts at least a portion of the double-stranded DNA to single-stranded DNA;
- b. hybridizing the single-stranded DNA with i) a first nucleic acid probe containing the point mutation at the 3' end and adapted with a moiety at the 5' end which can be captured to a solid support, and ii) a second nucleic acid probe labeled with a detectable moiety at the 3' end and which can hybridize at the 5' end with the single-stranded DNA adjacent the hybridized 3' end of the first nucleic acid probe;
- c. ligating the hybridized first and second nucleic acid probes;
- d. capturing the moiety on the first nucleic acid probe hybridized to the DNA to the solid support;
- e. denaturing the ligated first and second nucleic acid probes from the hybridized single-stranded DNA;
- f. removing uncaptured labeled probe; and,
- g. detecting the presence of captured detectable moiety, the presence of captured detectable moiety indicating the presence of the point mutation per at least about 75,000 normal copies in the sample.

## 12. A kit comprising:

- a) an exonuclease;
- b) a first nucleic acid probe which binds to target DNA and which is adapted with a capture moiety;
- c) a second nucleic acid probe which binds to target DNA adjacent the first probe and which is labeled with a detectable moiety; and
- d) a ligase.

Figure 1A. Zippered Gene-Specific Primer (zgsp)

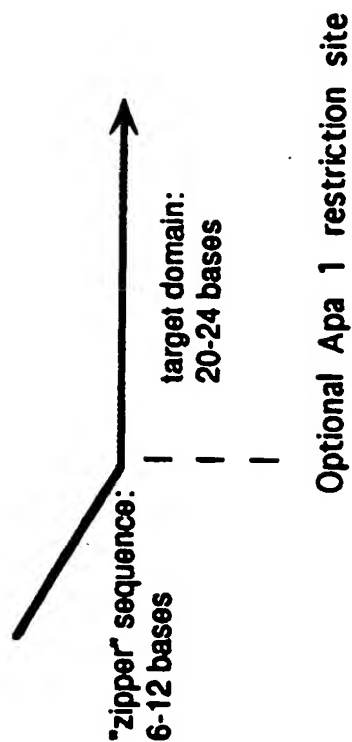
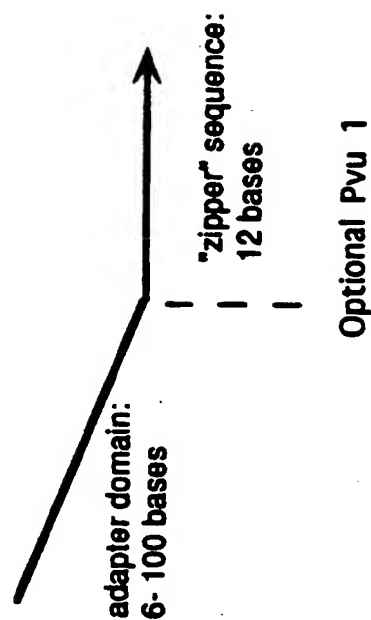


Figure 1B. Zipper Adapter Primer (zap)



2/17

Figure 2.

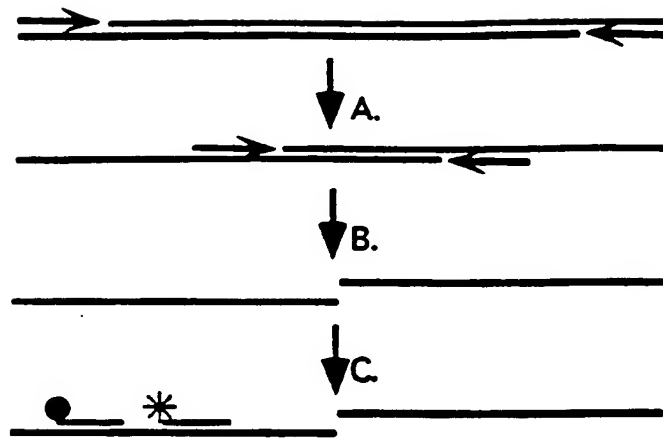
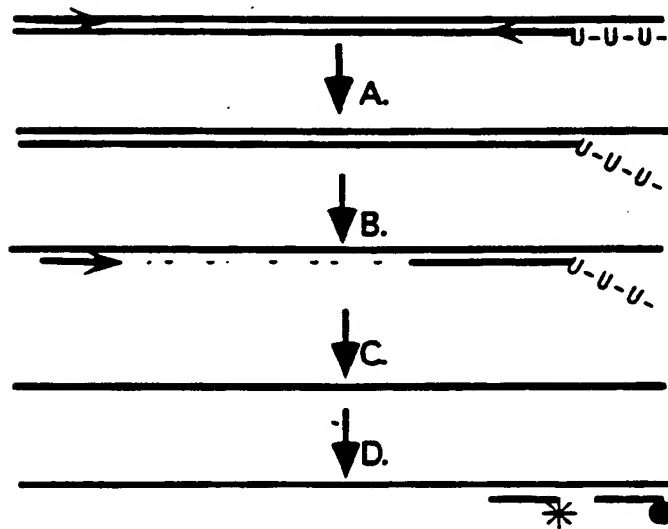
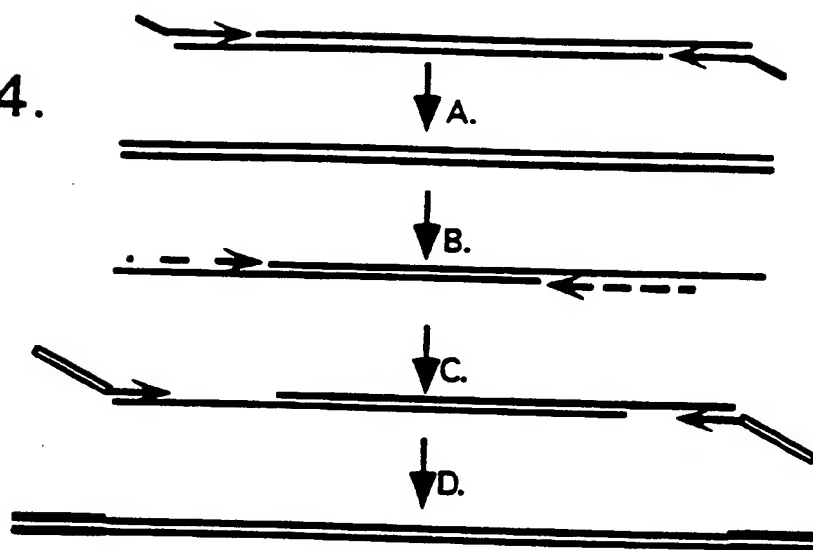


Figure 3.



3/17

Figure 4.



SUBSTITUTE SHEET (RULE 26)



4/17

Fig. 5A) Zap-F

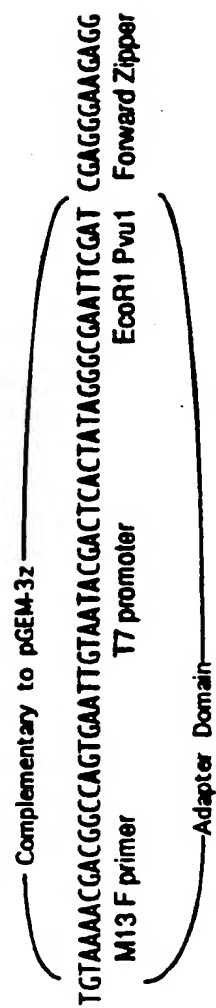
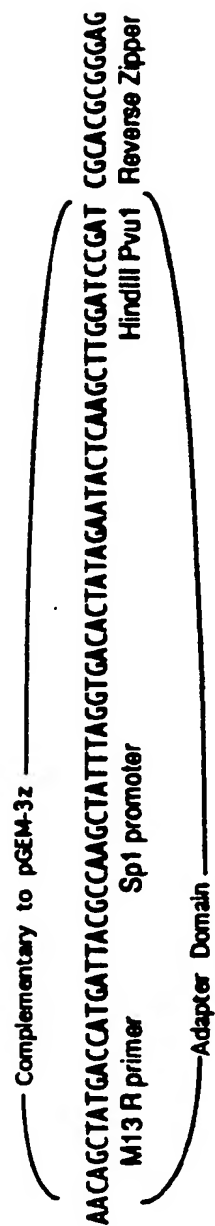


Fig. 5B) Zap-R



5/17

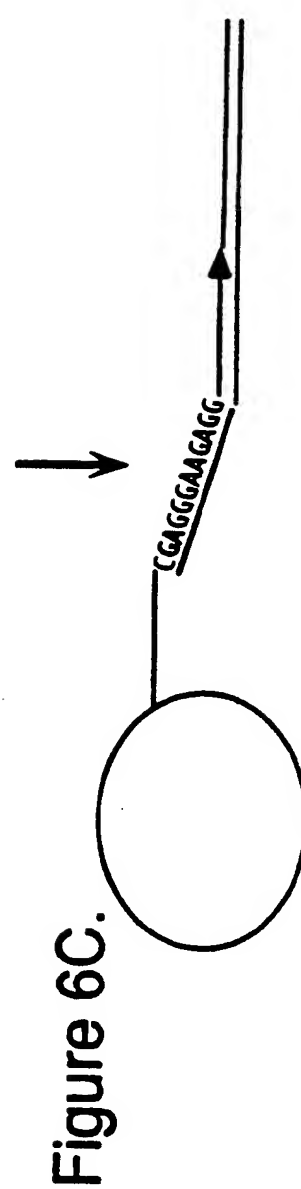
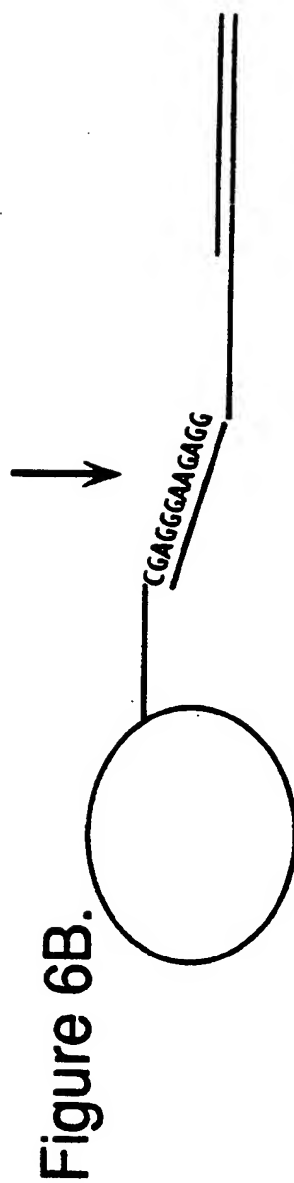
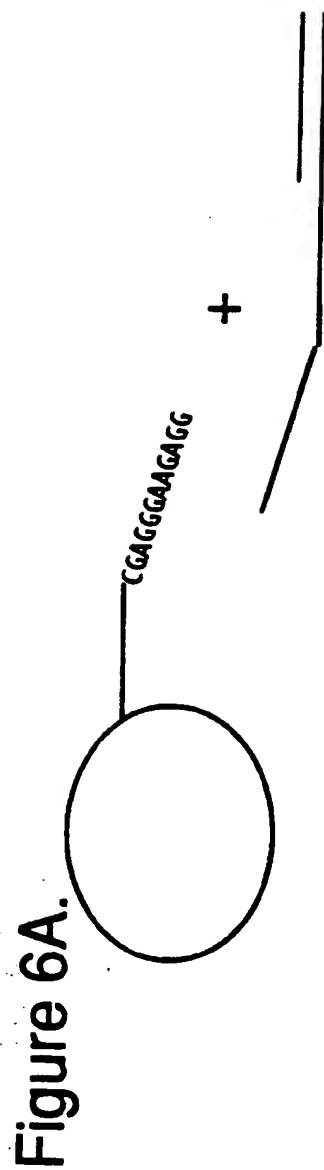
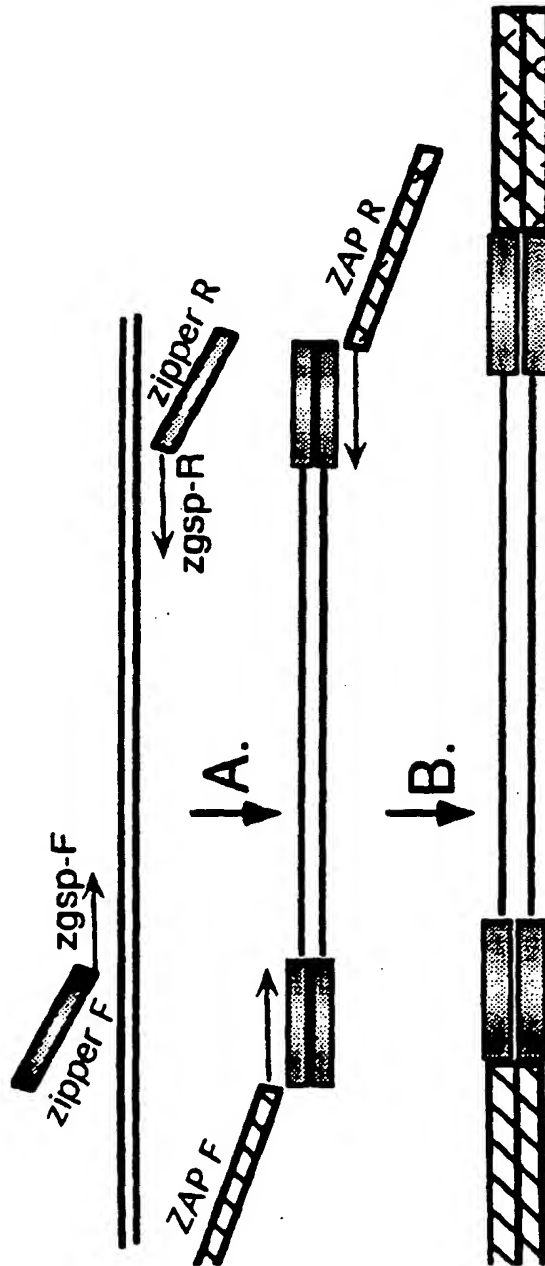
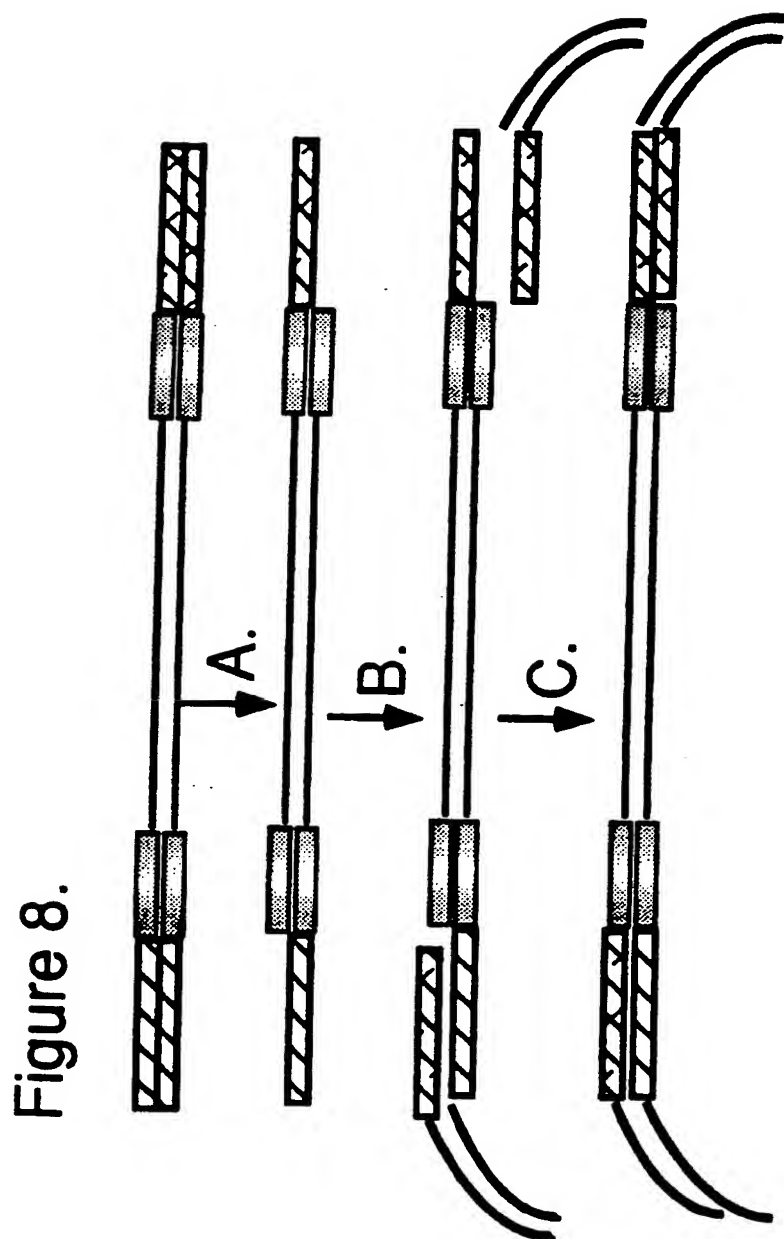


Figure 7.



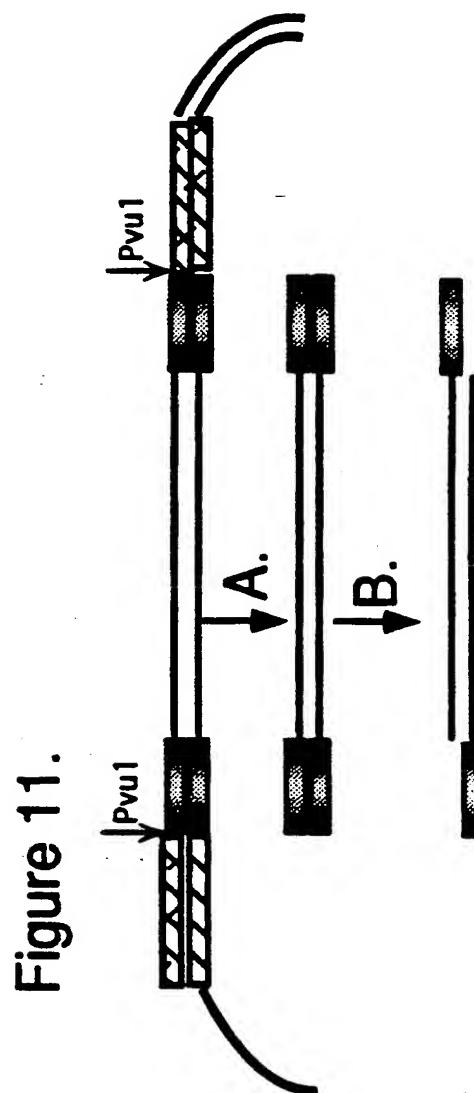
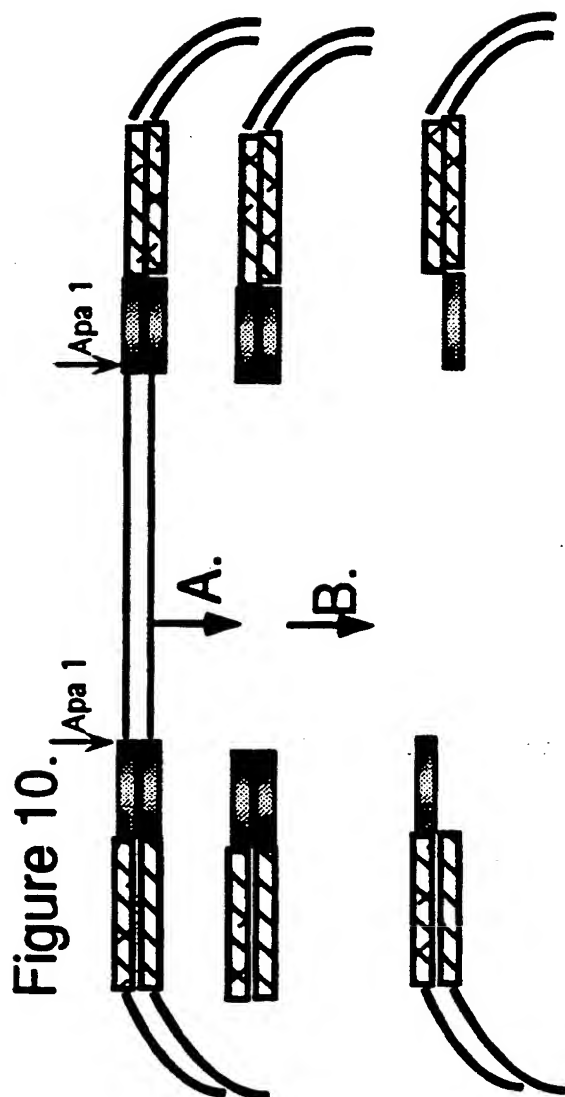


SUBSTITUTE SHEET (RULE 26)

8/17

**FIGURE 9.**

Bam H1	Apa1	Hind III
<u>GATCCATCGAGGGAAGAGGGGCCCTCTCTCCCGGAGAGGGCGCACGGTTCGA</u>		



10/17

Figure 12.

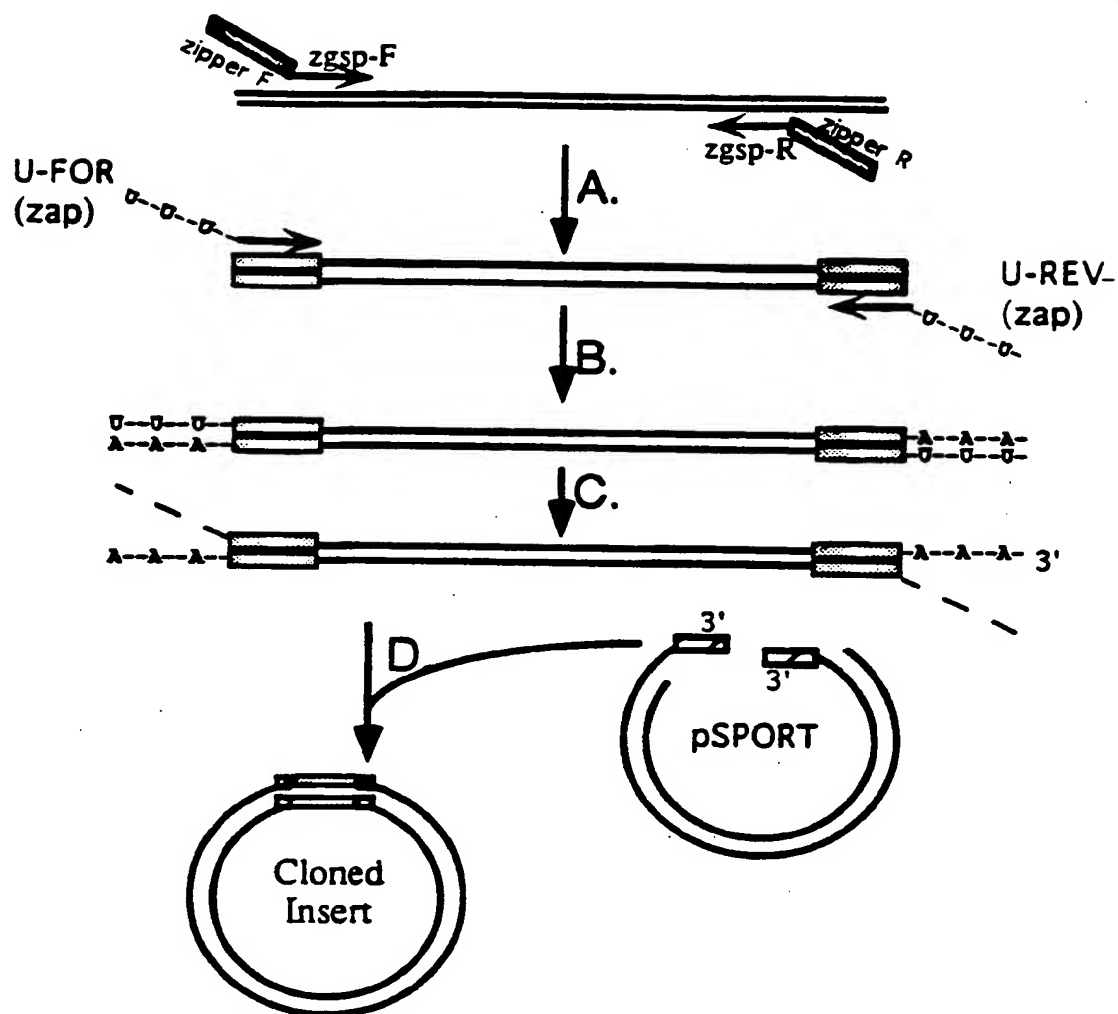
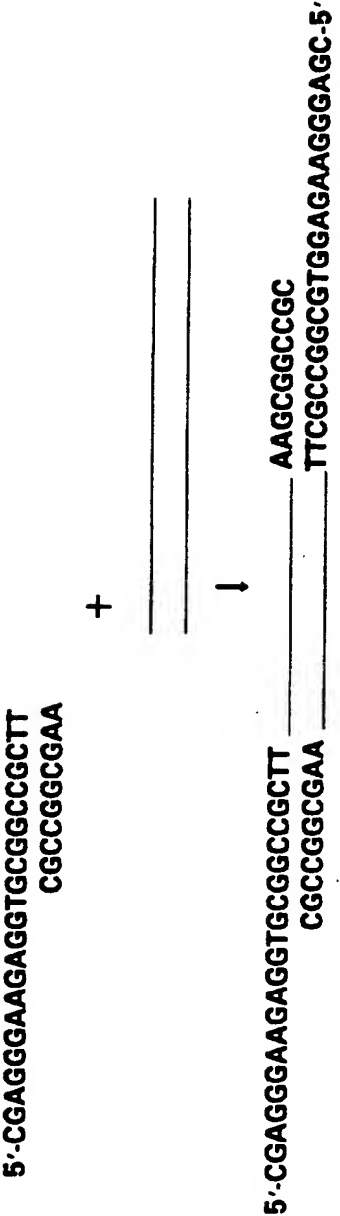


FIGURE 13A.





12/17

**FIGURE 13B.**

CGAGGGAAGAGGTGCGGCCGC	AATTGGGCGCGC
CGCCGGCGTTAA	CGCCGGCGTGGAGAAGGGAGC
CGAGGGAAGAGGTGCGGCCGC	GATCGCGGCCGC
CGCCGGCGCTAG	CGCCGGCGTGGAGAAGGGAGC
CGAGGGAAGAGGTGCGGCCGC	AGCTGCGGCCGC
CGCCGGCGTCGA	CGCCGGCGTGGAGAAGGGAGC

\*Note - same upper primer used in each case (blunt or sticky)  
CGAGGGAAGAGGTGCGGCCGC

Different lower primer used according to needs

GCGGCCGC	Blunt	zipper-adapter
AATTGGGCGCGC	EcoR I	zipper-adapter
GATCGCGGCCGC	BamH I	zipper-adapter
AGCTGCGGCCGC	Hind III	zipper-adapter

13/17

Figure 14.

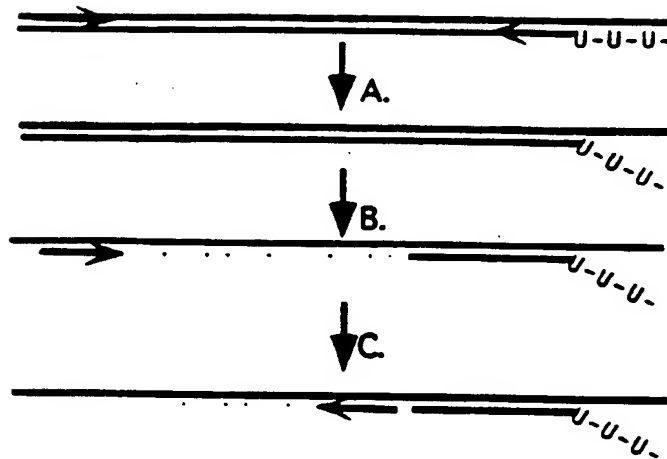
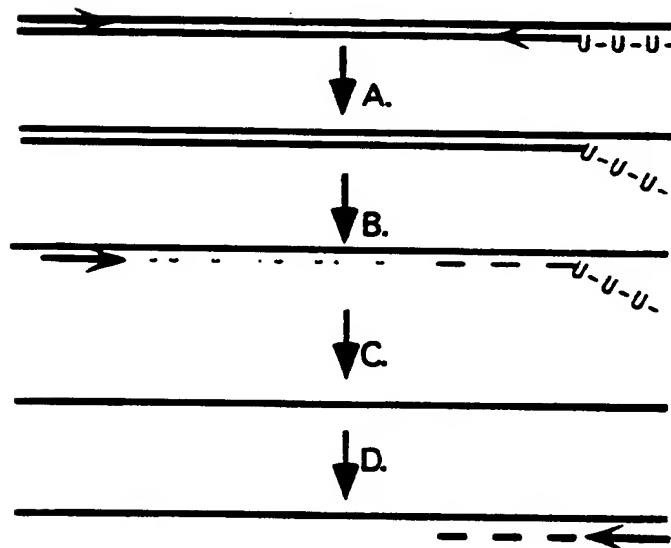


Figure 15.



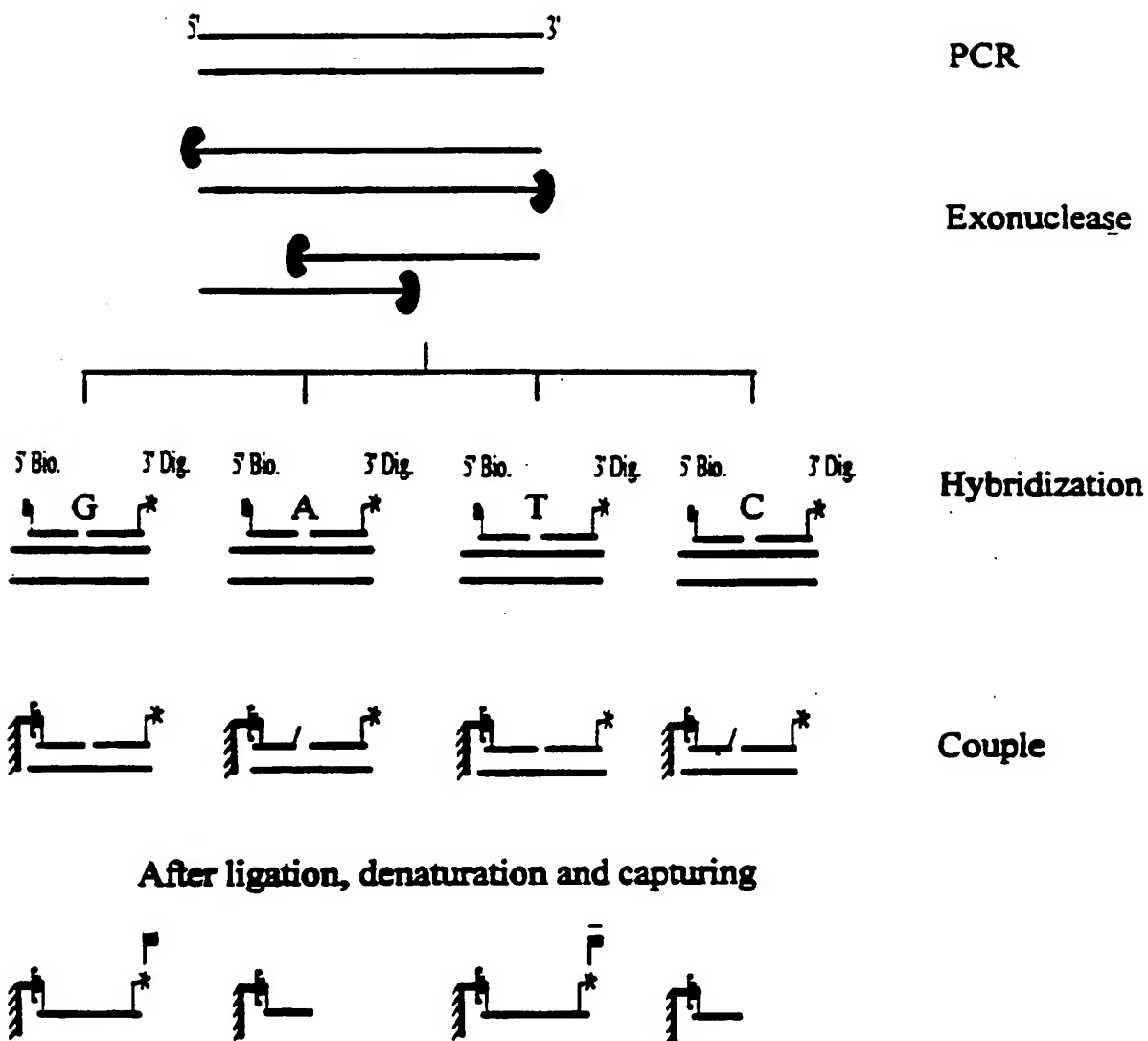
SUBSTITUTE SHEET (RULE 26)

14/17

Figure 16

Point-EXACCT

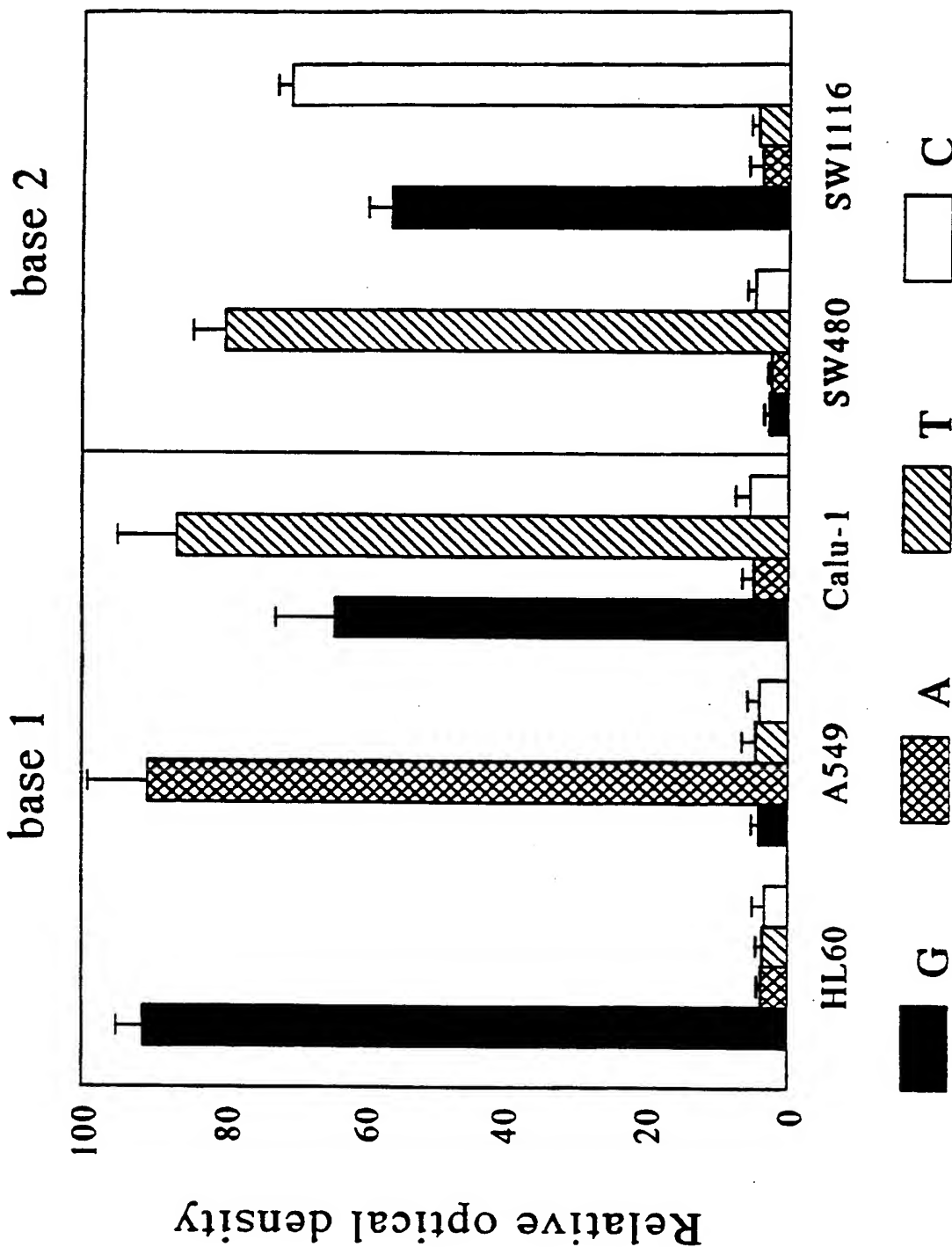
Calu-1



15/17

FIGURE 17

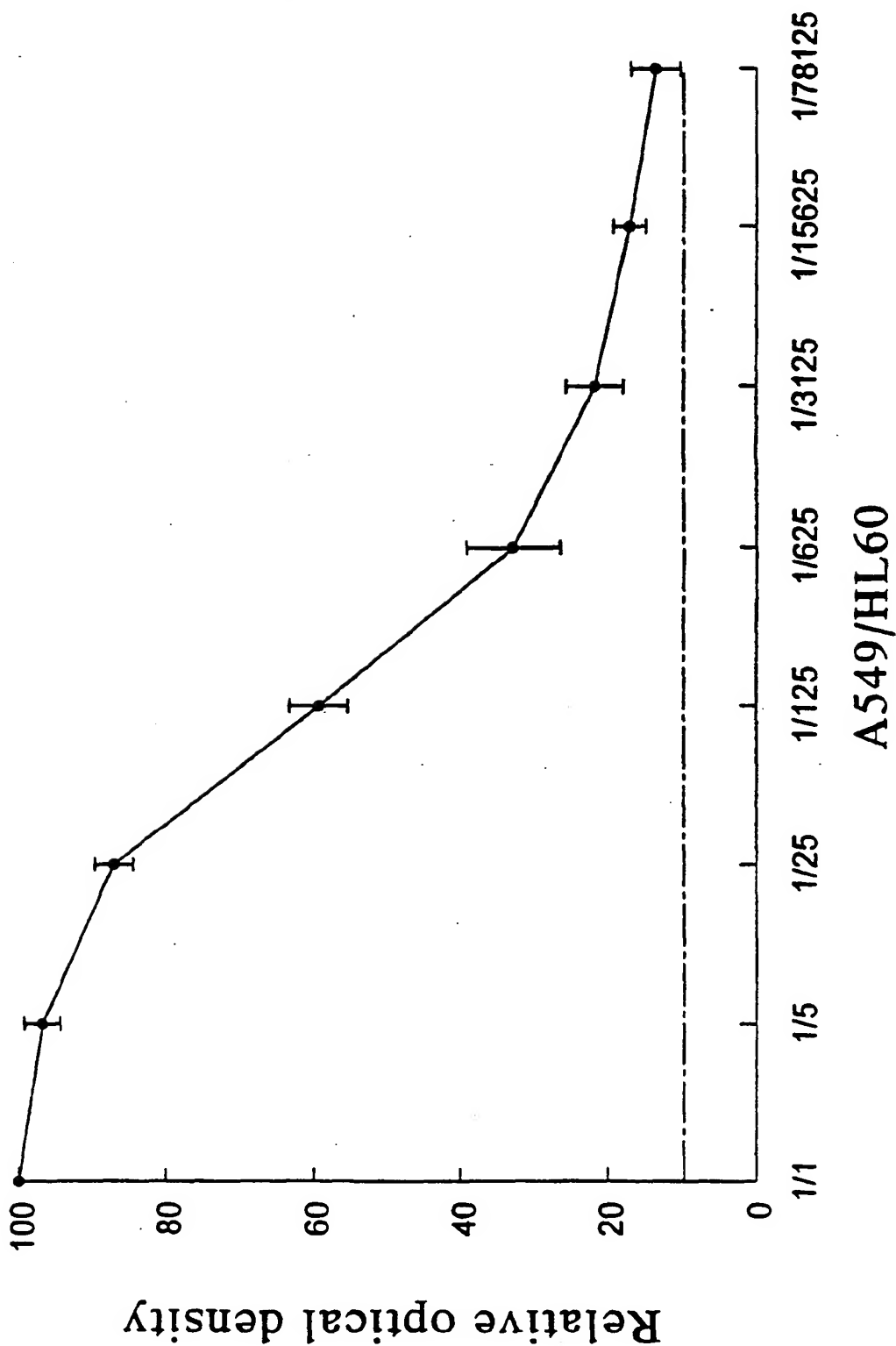
# Point-EXACCT



16/17

FIGURE 18A

# Reconstruction experiment a) A549/HL60

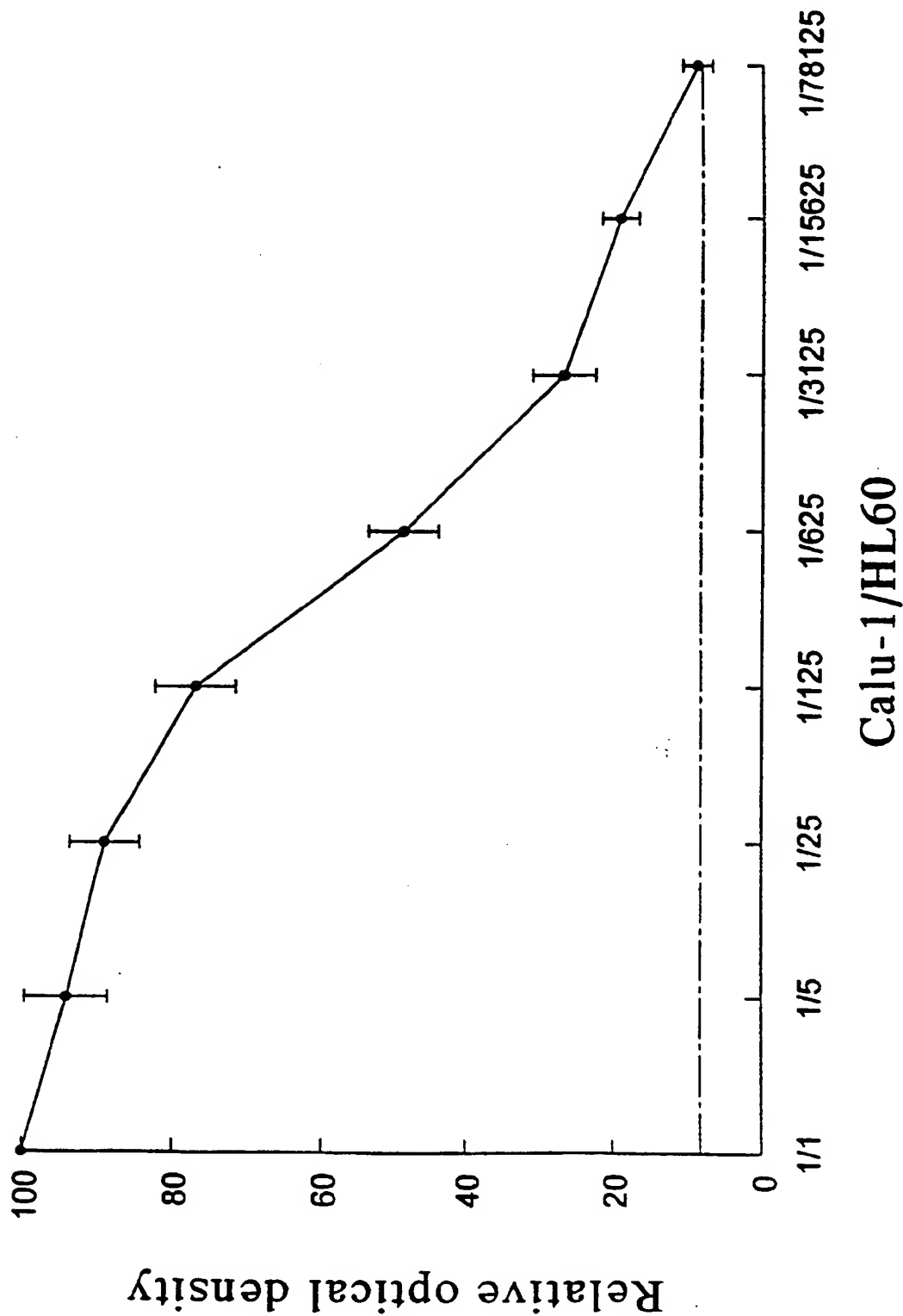


17/17

FIGURE 18B

# Reconstruction experiment

## b) Calu-1/HL60



## INTERNATIONAL SEARCH REPORT

Int. Application No

PCT/US 96/09208

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12Q1/68 C12P19/34 //C12N15/09,C12N15/10,C12N15/66,  
C07H21/04

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NUCLEIC ACIDS RESEARCH, vol. 22, no. 22, 11 November 1994, pages 4840-41, XP002013669 SOMERS V ET AL.: "A rapid, reliable method for detection of known point mutations : Point-EXACCT" see the whole document ---	1-12
Y	WO,A,94 24313 (MURTAGH) 27 October 1994 see page 16, line 26 - page 17, line 2; figures FIG,2 --- -/--	1-12

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

\*&\* document member of the same patent family

Date of the actual completion of the international search

18 September 1996

Date of mailing of the international search report

14. 10. 96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Osborne, H

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 96/09208

## C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>PROC. NATL. ACAD. SCI. USA, vol. 87, no. 22, November 1990, pages 8923-27, XP000209335 NICKERSON D ET AL: "Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay" see the whole document -----</p>	1-12



# INTERNATIONAL SEARCH REPORT

Inter-Application No  
PCT/US 96/09208

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9424313	27-10-94	US-A- 5518901	21-05-96
		AU-A- 6770994	08-11-94
		CA-A- 2161014	27-10-94
		EP-A- 0698121	28-02-96
-----			

